Problem Chosen	2021 MCM/ICM Summary Sheet	Team Control Number
F		2118464

A Study of National Higher Education System Based on Principal Component Analysis and Entropy Weight Method

Abstract

Higher education system is an important part of a country's efforts to provide its citizens with further education beyond primary and secondary education. Therefore, higher education system not only has the value of industry itself, but also has the value of training and educating citizens for the national economy.

For question one, this paper collects data of 15 indicators from 11 countries from 2016 to 2019 through various public websites, and then establishes an evaluation index selection model based on principal component analysis (PCA) to select the main influencing factors. Then, the health and sustainability index evaluation model of national higher education system based on entropy weight method (EWM) is established, and the health index evaluation model of national higher education system and the sustainability index evaluation model of national higher education system are obtained. Combined with the results of these two models, the health status of any national higher education system can be effectively evaluated.

For question two, this paper applies the model of question one to three countries, namely the United States, China and Vietnam. The final results are completely in line with the actual situation, which shows that the model has a certain reliability. According to the first mock exam result, the health evaluation index and the sustainable evaluation index of Vietnam's national higher education system are the lowest, and the lowest index is mainly due to their per capita GDP, higher education enrollment rate, and the proportion of public higher education funds to the government's expenditure. These main influencing indicators are all low, so this article chooses Vietnam as the country.

For question three, this paper selects three representative indicators after dimensionality reduction of the evaluation candidate indicators in question one, and puts forward a realizable and reasonable vision for Vietnam's education system.

For question four, after solving the model in question one, the health index of Vietnam's current national higher education system is 1106.49, and the sustainability index is 10.67. The health index of this system is 4072.47, and the sustainability index is 10.227. At this time, the health status of higher education system is close to that of China, which is a healthy and sustainable system.

For question five, this paper establishes a grey prediction model of policy time arrangement, adjusts the change parameters of key factors to affect the future data changes, and puts forward the targeted policy and implementation time planning for the transition from the current state to the proposed state.

For question six, the sensitivity test of the model proposed in this paper is carried out, and the current development characteristics and suggestions are summarized, and the suggestions and development situation are summarized in all aspects.

Keywords: higher education; principal component analysis; entropy weight method; a grey prediction model

Content

1 Restatement of the problem

1.1 Background

A system of higher education an essential factor in a nation's efforts to further educate its citizens beyond require primary as well as secondary education, and therefore enjoys value both as an industry itself and as a source of trained and educated citizens for the nation's economy. From Germany to The United States to Japan to Australia, numerous national approaches to higher education are proposed to not only educating students, but drawing large numbers of international students every year. Each of these systems of higher education has it strengths and limitations, and in the wake of adjustments required in the course of pandemic, nations have had the opportunity to reflect on what is working and what could be even better. Nevertheless, it's difficult to change, and the institutional changes require to advance any system require policies implemented over an extended period of time for the sake of reaching a more healthy and sustainable system.

1.2 Restatement of the Problem

In this problem, we should develop a model to measure and assess the health of a system of higher education at a national level, to identify a healthy and sustainable state for a give nation's higher education system. Meanwhile, we are supposed to propose and analyze a suite of policies to migrate a nation from its current state to your proposed healthy and sustainable state.

In addition, we are being asked to solve problems as followed:

- develop and validate a model or suite of models which could assess the health of any nation's system of higher education;
- apply our model to several countries, than selection a nation whose system of higher education has room for improvement according to our analysis;
- purpose an attainable and reasonable vision for our selected nation's system that supports a healthy and sustainable system of higher education;
- utilize our model to measure the health of both the current system and the healthy, sustainable system proposed for our selected nation;
- propose targets and available policies and an implementation timeline that will support the migration from the current state to our proposed state;
- use our model to shape and/or assess the effectiveness of our policies proposed above. And discuss the real-world impacts of implementing our plan both during the transition and in the end state, acknowledging the reality that change is hard.

2 Analysis of Problem

2.1 Analysis of Question 1

For the first problem, we collected relevant data, processed the data, and developed a model or a set of models that can evaluate the health status of higher education system in any country. This paper collected the relevant data of higher education in many countries to get the required data set, and finally completed the data collection. Because there were many indexes in quantitative analysis, it is easy to produce deviation. Therefore, this paper used principal component analysis to reduce and determined the dimension of indexes. The determination and evaluation method of index weight was also an important work. Meanwhile, the entropy weight method was used to build a set of national higher education system health and sustainable index evaluation model based on entropy weight method, to avoid the influence of subjectivity.

2.2 Analysis of Question 2

For the second problem, this paper applies the model of question one to the United States, China and Vietnam. According to the health and sustainability index of the national higher education system, it draws a conclusion and selects a country with room for improvement.

2.3 Analysis of Question 3

For the third problem, the representative indicators were selected after dimensionality reduction from the evaluation candidate indicators in question one, and a realizable as well as reasonable vision was put forward for the education systems of the United States, China and Vietnam selected in question two, so as to support a healthy and sustainable higher education system.

2.4 Analysis of Question 4

For the fourth problem, we used the model of question one to measure the health degree of the current system of the country selected in question two, and then utilized the model of question one to measure the health degree of the healthy and sustainable system of higher education described in question three, so as to get the result.

2.5 Analysis of Question 5

For the fifth problem, through the establishment of grey prediction of policy time scheduling model, by predicting the changes of key factors, the targeted policy and implementation schedule from the current state to the proposed state were proposed.

2.6 Analysis of Question 6

For the sixth problem, the effectiveness of your policy by summarizing all the models and discusses the real world impact of implementing the proposed measures during the transition period and in the final state.

3 Model Hypothesis

Regarding the establishment of the model, we have the following assumptions:

(1) The health evaluation models of national higher education system include health index evaluation model of national higher education system and sustainable index evaluation model of national higher education system.
(2) The health and sustainability index of national higher education system is independent of the four indicators of higher education foundation, higher education development performance, higher education investment and higher education process.
(3) The health status of the national higher education system is closely related to the health index and sustainability index of the national higher education system in the recent year, and the importance of each index is different, so the corresponding weight of each index is also different.
(4) If the health index and sustainability index of higher education system in two countries are the same, the health status of higher education system in two countries is the same.

4 Symbol Description

Symbol	Definitions
Per_GDP	GDP per capita
Edu_Enrollment	Enrollment rate of higher education
Pro_Researchers	Proportion of researchers by country
Phd_Research	Research on Doctoral Education Development Index
Num_University	Number of the top 500 international universities
Num_Subject	Number of subjects ranked in the top 100 internationally
Num_Nature	Number of Nature papers
Sci_Publication	The countries leading the world scientific publications
Cited_researcher	High number of cited researchers
Pro_Spengding	Poportion of public higher education spending in GDP
Pro_funds	Proportion of public higher education funds allocated by the government
Gdp_Spending	Research and development spending in universities as a percentage of GDP
T-S ratio	Teacher to student ratio
Youth-Rate	Youth literacy rate,ages 15-24
F-M ration	Female to male ratio,student at tertiary level education

5 Model Establishment and Solutions

5.1 Solution to Problem One

5.1.1 Scenario retelling

There are actual data in every indicator that assesses the health of the national higher education system. This article needs to use these data to determine the health and sustainability of a country's higher education system. Therefore, it is necessary to have an objective and accurate health evaluation index and sustainability evaluation index. The determination of the index needs to be fair and just, and fully consider the influence of each candidate on each index. Therefore, the team needs to establish a model to calculate the health evaluation index and sustainability evaluation index of the higher education system of different countries, so that the evaluation index can reasonably and fairly reflect the health status of the higher education system of each country.

5.1.2 Model I: Evaluation index selection model based on principal component analysis

The index evaluation data are published by international authoritative institutions and official statistical websites. The data mainly come from WIKI, Knoema, Our world in data, The world bank, World education news and reviews, so as to ensure the authority and consistency of the data.

Due to the global outbreak of the new crown epidemic in 2020, various industries have been affected to varying degrees. For the higher education industry, most countries will cancel offline teaching and switch to online teaching in 2020. Therefore, the higher education system data in 2020 cannot better reflect the health of the national higher education system.

The data collected in this article are from 2016 to 2019, ignoring the data in 2020. Data in four aspects are mainly collected, including higher education foundation, higher education development performance, higher education investment and higher education process. More indicators are extended below. Among them, the foundation of higher education and the performance of higher education development are the indicators of the national higher education system health index evaluation model. Higher education investment and higher education process are indicators of the sustainability index evaluation model of the national higher education system. Specific indicators are shown in Table 1.

Table 1: Specific indicators of the health and sustainability index evaluation model of the national higher education system

Evaluation Model of Health Index of Higher Education System	Foundation of Higher Education	Per_GDP
		Edu_Enrollment
		Pro_Researchers
		Phd_Research
	Higher education development	Num_University
		Num_Subject

	performance	Num_Nature Sci_Publication Cited_researcher
Sustainable Index Evaluation Model of Higher Education System	Higher education investment	Pro_Spengding Pro_funds Gdp_Spending
	Higher Education Process	T-S ratio Youth-Rate F-M ration

In order to select the evaluation indicators that affect the health status of the national higher education system, we used the principal component analysis method to calculate the information contribution rate of 9 candidate indicators for health evaluation and 6 candidate indicators for sustainable evaluation, and then selected Reasonable evaluation indicators.

1. Model establishment

（1） Standardized processing

Suppose the evaluation index variable has a total of m , which is recorded as $x_1, x_2, \ldots, x_m$ .And there are n national higher education systems to be evaluated, and the value of the j index of the i national higher education system is x_{ij}.Convert each index value x_{ij} into a standardized index x_{ij}, the conversion expression is as follows:

$$x_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j}, (i = 1,2,\cdots n; j = 1,2,\cdots,m)$$

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$$

$$S_j = \frac{1}{n-1}\sum_{i=1}^{n} x_{ij} - \bar{x}_j^2, (j = 1,2,\cdots,m)$$

Among them, x_j and S_j are the sample mean and standard deviation of the j index.

（2） Calculation of the correlation coefficient matrix R

The correlation coefficient matrix $\mathrm{R} = (r_{ij})_{m\times m}$ and r_{ij} are calculated as follows:

$$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$$

$$r_{ij}=\frac{\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)(x_{kj}-\bar{x}_j)}{\sqrt{\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)^2\sum_{k=1}^{n}(x_{kj}-\bar{x}_j)^2}},(i=1,2,\dots n;\ j=1,2,\dots m)$$

Where $\bar{x}_i$ is the sample mean of the i index, and $r_{ii}=1$, $r_{ij}=r_{ji}$, and r_{ij} are the correlation coefficients between the i index and the j index.

(3) Calculate eigenvalues and eigenvectors

Next, calculate the eigenvalue $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$ of the correlation coefficient matrix R and the corresponding eigenvector $\mu_1,\mu_2,\cdots,\mu_m$, where $\mu_j=(\mu_{1j},\mu_{2j},\cdots,\mu_{nj})^T$ is composed of eigenvectors and m new index variables are:

$$\begin{aligned}
y_1 &= \mu_{11}x_1+\mu_{21}x_2+\cdots+\mu_{n1}x_n\\
y_2 &= \mu_{12}x_1+\mu_{22}x_2+\cdots+\mu_{n2}x_n\\
&\cdots\\
y_m &= \mu_{1m}x_1+\mu_{2m}x_2+\cdots+\mu_{nm}x_n
\end{aligned}$$

In the expression, y_1 is the first principal component, y_2 is the second principal component, and so on, y_m is the m-th principal component.

(4) Calculate information contribution rate and cumulative contribution rate

First, the information contribution rate and cumulative contribution rate of the feature value $\lambda_j(j=1,2,\cdots,m)$ need to be calculated. The information contribution rate b_j of the principal component y_j is calculated as follows:

$$b_j = \frac{\lambda_j}{\sum_{k=1}^{m}\lambda_k}\ (j = 1,2,\cdots, m)$$

Where b_j is the information contribution rate of the j principal component.

The cumulative contribution rate of the principal components $y_1,y_2,\cdots y_p$ is calculated as follows:

$$\alpha_p=\frac{\sum_{k=1}^{p}\lambda_k}{\sum_{k=1}^{m}\lambda_k}$$

When α_p is close to 1, we will select the first p index variables $y_1,y_2,\cdots y_p$ as p principal components to replace the original m evaluation index variables.

2. Model solution

First, the data is standardized, and then the correlation coefficient between each evaluation index is calculated to form a correlation coefficient matrix. The correlation coefficient matrix can be used to calculate the eigenvectors shown in Table 2 and Table 3.

Table 2: The feature vector corresponding to the principal components of the health index evaluation model of the higher education system

	Per_GDP	Edu_Enrollment	Pro_Researchers	Phd_Research	Num_University	Num_Subject	Num_Nature	Sci_Publication	Cited_researcher
characteristic vecter 1	37.312	-26.317	4.426	.692	-14.479	-5.651	-27.670	19.814	21.174
characteristic vecter 2	-26.317	21.514	-3.523	-1.893	14.150	-5.112	14.545	-11.391	-7.741
characteristic vecter 3	4.426	-3.523	47.634	-7.445	26.532	-37.028	-30.446	-38.190	40.708
characteristic vecter 4	.692	-1.893	-7.445	5.049	-12.665	17.136	2.616	8.464	-10.265
characteristic vecter 5	-14.479	14.150	26.532	-12.665	54.341	-76.892	-14.565	-31.699	50.421
characteristic vecter 6	-5.651	-5.112	-37.028	17.136	-76.892	236.897	47.325	36.347	-210.083
characteristic vecter 7	-27.670	14.545	-30.446	2.616	-14.565	47.325	77.832	-11.486	-72.150
characteristic vecter 8	19.814	-11.391	-38.190	8.464	-31.699	36.347	-11.486	59.012	-22.636
characteristic vecter 9	21.174	-7.741	40.708	-10.265	50.421	-210.083	-72.150	-22.636	216.038

Table 3: The characteristic vector corresponding to the principal components of the sustainability index evaluation model of the higher education system

	Pro_Spengding	Pro_funds	Gdp_Spending	T-S ratio	Youth-Rate	F-M ration
characteristic vecter 1	2.479	-1.341	.343	-.750	-.460	-.359
characteristic vecter 2	-1.341	2.988	.297	-.748	-.998	-.859
characteristic vecter 3	.343	.297	2.258	-2.095	-1.048	.007
characteristic vecter 4	-.750	-.748	-2.095	3.576	1.936	.091
characteristic vecter 5	-.460	-.998	-1.048	1.936	2.261	.392
characteristic vecter 6	-.359	-.859	.007	.091	.392	1.659

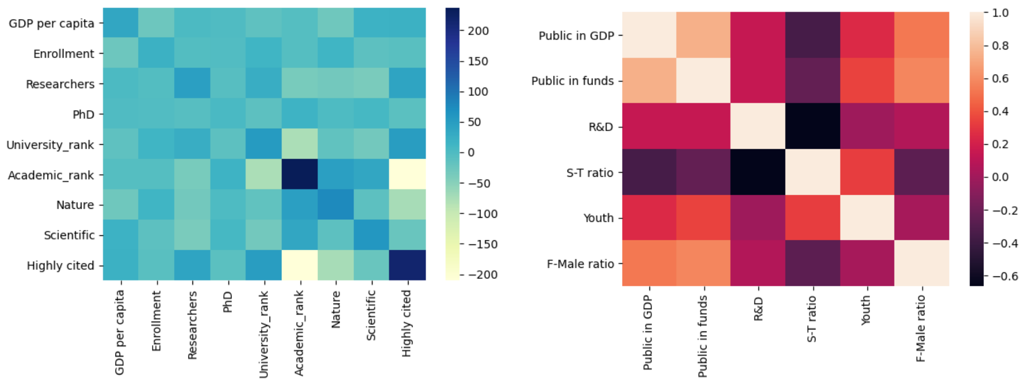


Figure 1: Feature vector heat map

Using eigenvectors and standardized data, the corresponding principal components can be obtained, and the contribution rate and cumulative contribution rate of each principal component can be calculated using the eigenvalues of the correlation coefficient matrix. The results are shown in Table 4.

Table 4: Principal component analysis results

	Initial Eigenvalues			Extraction Sums of Squared ..	
Component	Total	% of Variance	Cumulative %	Total	% of Variance
1	7.399	82.216	82.216	7.399	82.216
2	1.041	11.567	93.782	1.041	11.567
3	.550	6.116	99.898		
4	.008	.088	99.986		
5	.001	.011	99.997		
6	.000	.002	99.999		
7	3.804E-5	.000	100.000		
8	1.810E-5	.000	100.000		
9	1.924E-8	2.138E-7	100.000		

	Initial Eigenvalues			Extraction Sums of Squared ..	
Component	Total	% of Variance	Cumulative %	Total	% of Variance
1	4.328	72.132	72.132	4.328	72.132
2	1.485	24.744	96.876	1.485	24.744
3	.186	3.105	99.981		
4	.001	.017	99.998		
5	.000	.002	100.000		
6	1.444E-7	2.407E-6	100.000		

According to the principal component analysis results in Table 4, the cumulative contribution rate of the first four principal components of the health index evaluation model of the higher education system exceeds 99.9%. Therefore, we selected the first four principal components as the indicators for health evaluation; The cumulative contribution rate of the first four principal components of the education system sustainability index evaluation model exceeds 99.9%. Therefore, we selected the first four principal components as indicators for sustainability evaluation.

5.1.3 Model II: National Higher Education System Health and Sustainability Index Evaluation Model Based on Entropy Method

Entropy was first introduced into information theory by Shennong, and it has been widely used in engineering technology, social economy and other fields. The basic idea of the entropy method is to determine the objective weight according to the variability of the index.

The entropy method uses information entropy to evaluate the degree of difference of an indicator in different evaluation objects. If the data of an indicator changes greatly, it means that the amount of information provided by the indicator is large, the corresponding information entropy is small, and the weight assigned is large; vice versa. If the data of a certain index are all equal, it means that the index cannot provide distinguishing information to the evaluation object, and the corresponding information entropy is the largest, and the weight is zero. Therefore, the entropy method can be used to objectively and reasonably obtain the weight of each indicator, and then obtain the national higher education system health and sustainability index evaluation model[1].

1. **Model establishment**

（1） Data normalization

The data of n indicators are normalized. Suppose that q countries $X_1, X_2, \ldots, X_k$ are given, where

$$X_i = \{x_1, x_2, \ldots, x_n\}$$

Assuming that the normalized value of each indicator data is $Y_1, Y_2, \ldots, Y_k$

Then $Y_{ij} = \dfrac{X_{ij} - \min(X_i)}{\max(X_i) - \min(X_i)}$, where i is the i country and j is the j indicator.

（2） Find the information entropy of each index

According to the definition of information entropy in information theory, the information entropy of a set of data is:

$$E_j = -\frac{1}{\ln n}\sum_{i=1}^{n} p_{ij} \ln p_{ij}$$

Among them, $p_{ij} = \frac{Y_{ij}}{\sum_{i=1}^{n} Y_{ij}}$, if $p_{ij} = 0$, then define $\lim_{p_{ij}\to 0} p_{ij} \ln p_{ij} = 0$

（3） Determine the weight of each index

According to the calculation formula of information entropy, the information entropy of each indicator is calculated as $E_1, E_2, \ldots, E_k$. The weight of each indicator is calculated by information entropy:

$$W_i = \frac{1 - E_i}{k - \sum E_i} (i = 1, 2, \ldots, k)$$

（4） The calculation of the health evaluation index of the national higher education system

After obtaining the weight of each indicator, the evaluation index can be calculated:

$$Grade_k = \sum_{j=1}^{n} W_j X_{kj}$$

2. Model solution

（1） Data normalization

According to the established national higher education system health and sustainability index evaluation model based on the entropy method, we can get the normalized results of each indicator data, and the obtained data are shown in the following table:

Table 5: Normalized results of various indicator data for some countries

Country	Per_GDP	Edu_Enrollment	Pro_Researchers	Phd_Research
China	0	0.005156242	0.017297774	0.663524898
India	0.003131575	0	0	0
Vietnam	0.001181806	0.024966424	0.064088151	0.902776139
Brazil	0.212880985	0.859068213	1	0.91201641
America	0.068771596	0.076546137	0.45357646	0.889265716
Germany	0.002365486	0.023711404	0.087624604	0.997583107
British	0.000134863	0.016032892	0.060801338	0.885110019
France	8.4431E-06	0.015299888	0.058072558	0.872455364
Italy	0.00118395	0.023043071	0.084492194	1
Canada	1	1	0.826968582	0.005930743
Japan	0.197285212	0.223720045	0.952589299	0.46595769

Table 6: The normalized results of each indicator data of some countries

Country	Pro_Spengding	Pro_funds	Gdp_Spending	T-S ratio
China	0.231886576	0.094785139	0.65973305	0.734118062
India	0	0	0	0.176497284
Vietnam	0.328656125	0.146425651	0.890757929	0.923663977
Brazil	1	1	1	0.673913211
America	0.334986338	0.179181946	0.872769154	0.879089733
Germany	0.353361653	0.148464551	0.978529085	0.996829948
British	0.312137691	0.129786232	0.871063807	0.909361193
France	0.307518019	0.127714188	0.859039121	0.899443379
Italy	0.354243213	0.148456356	0.980935427	1
Canada	0.544982544	0.895777482	0.187587518	0
Japan	0.175816222	0.122682863	0.49077902	0.421814955

In order to solve the problem that the different dimensions of various indicators cannot be directly summarized, the processed data is subjected to a dimensionless operation-normalization, which makes the data more stable without much fluctuation, which is convenient for calculating information entropy. The following is a comparison chart of a certain index of the two models before and after normalization. It can be seen that there is a significant difference between the index values before and after normalization.

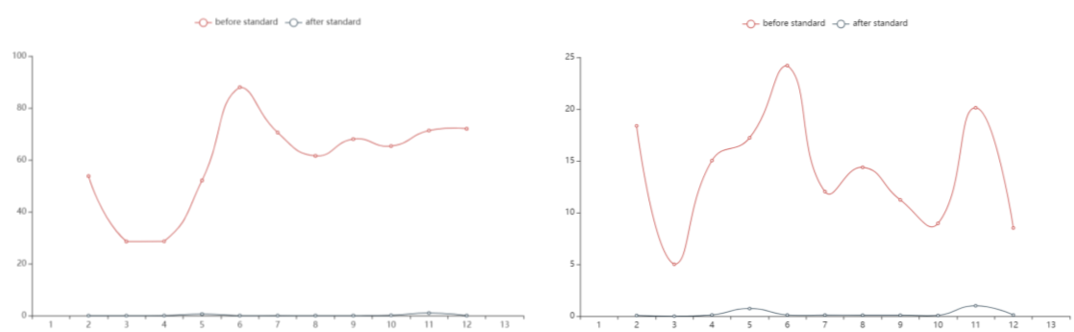


Figure 2: Comparison of two indicators before and after normalization

According to the calculation formula of information entropy, the information entropy of each of the eight indicators of the two models can be calculated as follows:

Table 7: Information entropy table of 4 indicators of the health index evaluation model of the higher education system

	Per_GDP	Edu_Enrollment	Pro_Researchers	Phd_Research
Information entropy	0.4832	0.5883	0.9351	0.6857

Table 8: Information entropy table of 4 indicators of sustainability index evaluation model of higher education system

	Pro_Spengding	Pro_funds	Gdp_Spending	T-S ratio
Information entropy	0.9073	0.7725	0.9338	0.9307

After obtaining the information entropy of each indicator, according to the calculation formula of the indicator weight, the weight of each indicator can be obtained as shown in the following table:

Table 9: Weight table of the 4 indicators of the health index evaluation model of the higher education system

	W_1	W_2	W_3	W_4
Weights	0.3952	0.3148	0.0496	0.2403

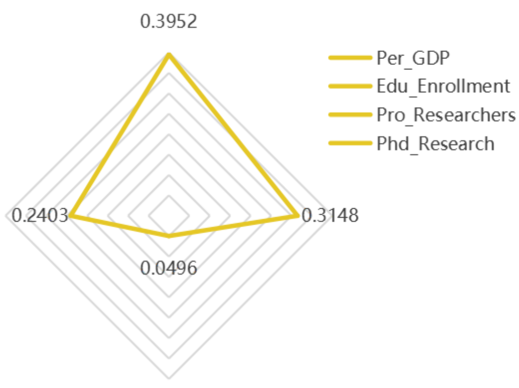


Figure 3: Radar chart of the weights of the 4 indicators of the health index evaluation model of the higher education system

Table 10: Weight table of the 4 indicators of the sustainability index evaluation model of the higher education system

	W_1	W_2	W_3	W_4
Weights	0.2034	0.4992	0.1453	0.1521

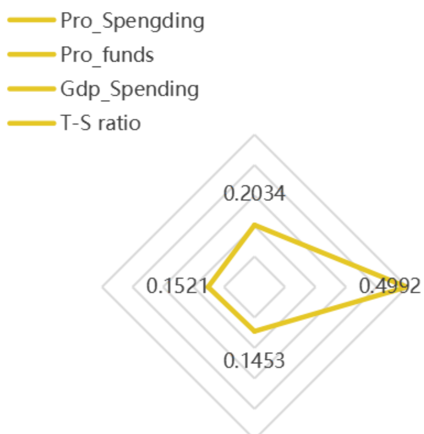


Figure 4: Radar chart of the weights of the four indicators of the sustainability index evaluation model of the higher education system

The larger the evaluation index data calculated at the end, the better the health or sustainability.

5.2 Solution to Problem Two

5.2.1 Scene retelling

This paper applies the model of question one to the United States, China and Vietnam, and draws a conclusion to analyze and select a country with room for improvement in higher education system.

5.2.2 Solution of problem two

According to the collected data sets, the eight indicators of the United States, China and Vietnam are shown in Table 11

Table 11: Data table of 8 indicators for 3 countries

	Per_GDP	**Edu_Enrollment**	**Pro_Researchers**	**Phd_Research**
America	65281	88.03	138.3	3

China	10262	53.76	213.4	1
Vietnam	2777	28.64	0	0
	Pro_Spengding	**Pro_funds**	**Gdp_Spending**	**T-S ratio**
America	7.05	24.21	2.834	0.083
China	3.81	18.38	2.212	0.041
Vietnam	4.78	15.03	0.498	0.048

According to the weight obtained from the solution of question 1, the health evaluation indexes of the national higher education system of the United States, China and Vietnam are 25834.34, 4083.29 and 1106.49 respectively, and the sustainable evaluation indexes of the national higher education system are 13.94, 10.28 and 8.56 respectively. According to the value of each weight obtained from question 1, the main factors affecting the health of the national higher education system are the per capita GDP and the enrollment rate of higher education, and the sustainable situation of the national higher education system is the proportion of public higher education funds to the government funds. Vietnam's health evaluation index and sustainability evaluation index are low, mainly due to its low per capita GDP, enrollment rate of higher education, and the proportion of public higher education funds in government funds.

5.3 Solution to Problem Three

According to the model constructed in question 1, this paper selects three representative indicators from nine health evaluation candidate indicators and six sustainability evaluation candidate indicators after dimensionality reduction, and their weights in the evaluation model are all above 0.3, which are: per capita GDP, higher education enrollment rate, and the proportion of public higher education funds in government expenditure. In this regard, this paper puts forward a realizable and reasonable vision for the education systems of the United States, China and Vietnam selected in question 2, so as to support a healthy and sustainable higher education system.

5.3.1 Advice for Per_GDP

This is the basic condition for the people's life. Only when the living material is rich can the higher education system have a better space to play. For Vietnam, at the national level, it can do a good job in infrastructure construction and develop the primary, secondary and tertiary industries; at the enterprise level, it can expand investment and actively attract investment; at the individual level, it can improve the level of legal income and improve the living standard.

5.3.2 Advice for Edu_Enrollment

The greatest contribution of higher education to the society lies in the cultivation of talents, who can promote social development for the society[3]. The core of higher education is personnel training. For Vietnam, in order to improve the enrollment rate of higher education, we should strengthen the construction of basic education and secondary education system.

5.3.3 Advice for Pro_funds

The types of R & D activities are divided into three categories: basic research, applied research and experimental development. Among them, basic research embodies the original innovation ability facing the scientific frontier[4]. The higher the proportion of R & D funds in government expenditure, the more attention should be paid to basic research in Colleges and universities. Improving this proportion also helps to improve the innovation enthusiasm of scientific research personnel. For Vietnam, increasing the proportion of R & D funds in Colleges and universities is conducive to the construction of basic research and improving the potential of applied research and experimental development.

Generally speaking, Vietnam's higher education system is basically unable to reach the health status of the higher education system of the United States, but it can keep up with China and strive to achieve the health and sustainable status of China's higher education system.

5.4 Solution to Problem Four

In question two, this paper identifies Vietnam as a country with room for improvement in its higher education system. In the third question, this paper puts forward suggestions from three aspects: GDP per capita, higher education enrolment rate, and the proportion of public higher education funds to government funds, and expects Vietnam to achieve the healthy and sustainable status of China's current higher education system.

We use a model of the current higher education system of Vietnam and proposed in this paper are applied to solve the healthy and sustainable system in question 3 hope that Vietnam's higher education system in this paper can in per capita GDP, higher education enrollment of higher education funds, public government points out that the proportion of funding from three aspects and the current situation of China, as a result, eight indicators data as shown in table 12.

Table 12: Indicators of the current and proposed higher education system in Vietnam

	Per_GDP	Edu_Enrollment	Pro_Researchers	Phd_Research
Now_Vietnam	2777	28.64	0	0
After_Vietnam	10262	53.76	0	0
	Pro_Spengding	Pro_funds	Gdp_Spending	T-S ratio
Now_Vietnam	4.78	15.03	0.498	0.048

After_Vietnam	4.78	18.38	0.498	0.048

After the solution of Model II, the health index of Now_Vietnam is 1106.49, and the sustainability index is 10.67. After_Vietnam health index is 4072.47, and its sustainability index is 10.227. At this time, the health status of higher education system is similar to that of China, and it is a healthy and sustainable system.

5.5 Solution to Problem Five

5.5.1 Model III: Policy Timing Model Based on Grey Forecasting

Based on people's understanding of the uncertainty characteristics of system evolution, grey prediction uses sequence operators to generate and process the original data, mine the law of system evolution, establish the grey system model, and make scientific quantitative prediction of the future state of the system.

In response, this paper collected data on Vietnam from 2016 to 2019 from the three key factors selected in Question 3.

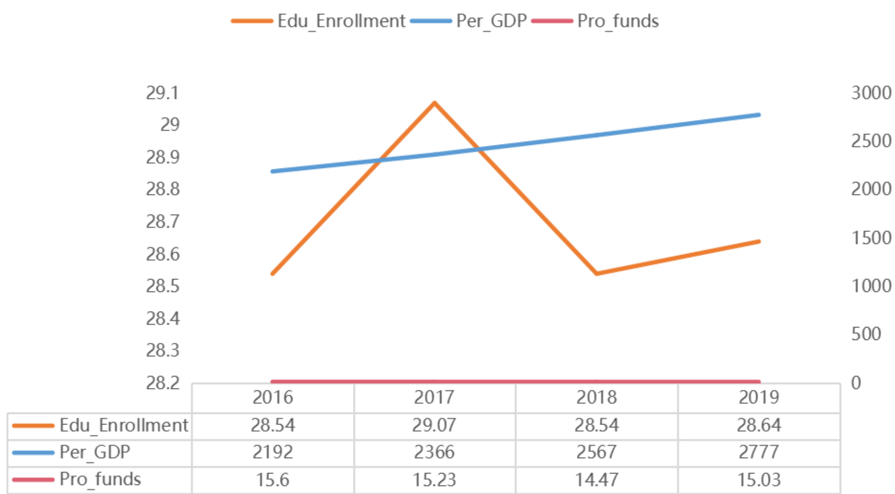


Figure 4: Key factor data for Vietnam

First, establish the time series:

$$x^{(0)} = (x^0(1), x^0(2), \dots x^0(n))$$

Find the order ratio L(k) of the above equation:

$$L(k) = \frac{x^0(k-1)}{x^0(k)}$$

After inspection, all of the L(k) in $[e^{-\frac{2}{n+1}}, e^{\frac{2}{n+1}}]$ range, can be satisfied the GM (1, 1) model.

GM(1,1) model was established with $x^{(0)}$ as the data column:

$$x^{(0)}(k) + az^{(1)}(k) = b$$

Estimated values of a and b were obtained by regression analysis, so the corresponding bleaching model is:

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b$$

The solution of the above equation is:

$$x^{(1)}(t) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(t-1)} + \frac{b}{a}$$

The predicted value is thus obtained:

$$x^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a}, k = 1,2,\ldots,n-1$$

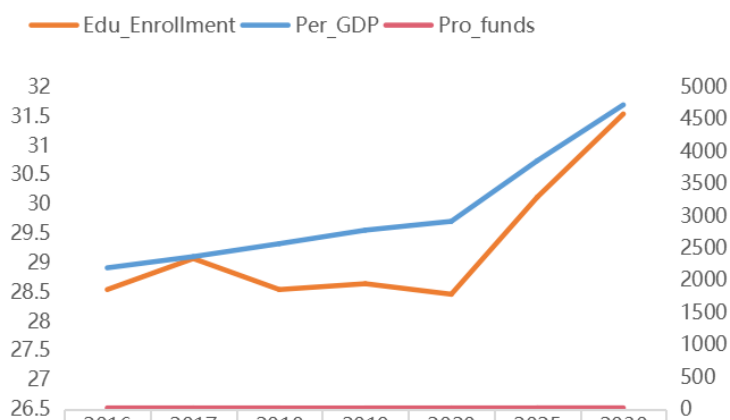
Thus, the predicted value is obtained accordingly:

$$x^{(0)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k), k = 1,2,\ldots,n-1$$

After calculating the result, we carried out residual test on the predicted value:

$$\varepsilon(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)}, k = 1,2,\ldots,n$$

After inspection, all $|\varepsilon(k)| < 0.1$, think reached higher requirements.



	2016	2017	2018	2019	2020	2025	2030
Edu_Enrollment	28.54	29.07	28.54	28.64	28.46	30.12	31.54
Per_GDP	2192	2366	2567	2777	2914	3856	4723
Pro_funds	15.6	15.23	14.47	15.03	15.45	16.57	17.43

Figure 5: Key factor data and policy projections for Vietnam

Based on the grey prediction results and the policy recommendations in Question 3, this paper proposes targeted policies and implementation timetables for migrating from the current state to the suggested state:

Policies for GDP per capita:

For Vietnam, economic activity should revive manufacturing and service industries in 2021, and the job market should provide more non-agricultural jobs to ease the pressure of social unemployment; We should continue to do a good job in infrastructure construction and develop the primary, secondary and tertiary industries for peaceful development. The government should continue to strengthen support for the poor and strengthen social infrastructure construction from now on.

Policies for higher education enrolment:

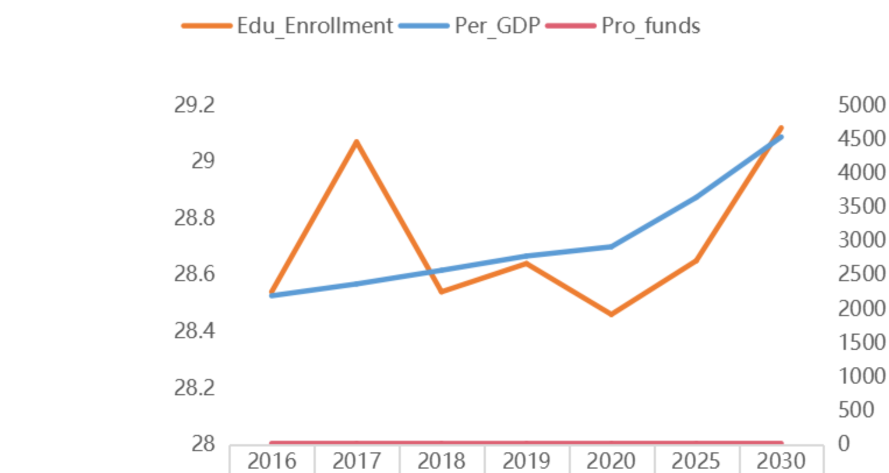
For Vietnam, policies related to basic education should be formulated in 2022, and the preliminary plan should be realized in 2025, so as to strengthen the learning and training of children and comprehensively improve the social and humanistic quality so as to increase the enrollment rate of higher education [5].

Policies for public higher education as a percentage of government spending:

For Vietnam, policies should be formulated in 2022 to increase the proportion of research and development funds in the GDP of universities, and continue to strengthen the welfare subsidies and technical support for researchers.

5.6 Solution to Problem Six

For the model in Question 5, after removing the influence of key factors brought by policies, the following data can be obtained through grey prediction:



	2016	2017	2018	2019	2020	2025	2030
Edu_Enrollment	28.54	29.07	28.54	28.64	28.46	28.65	29.12
Per_GDP	2192	2366	2567	2777	2914	3642	4532
Pro_funds	15.6	15.23	14.47	15.03	15.45	15.48	15.73

Figure 6: Key factor data and forecasts without policy in Vietnam

From the chart above, we can see that effective policies bring positive benefits, and that implementing them during the transition and in the final state has some impact on the real world:

For countries, increasing the proportion of higher education will help develop a better higher education system. Higher education serves science and technology and promotes the rapid development of society. Higher education cultivates talents, provides human resources for social development, and becomes a powerful driving force for social sustainable development.

For schools and teachers, the knowledge superiority and talent superiority of teachers in higher education are easy to play the role of the main force of high-tech research. Favoring policies for teachers will help high-tech talents give full play to their research advantages and make breakthroughs in basic fields. It is beneficial to improve the efficiency of education, reduce the cost of education, introduce high-level teachers and improve the quality of education.

For students, higher education stimulates and develops habits of thought, imparts knowledge to students, teaches them personal responsibilities to society, the country and the world, and imparts skills to make a living.

However, from the current situation, it is still difficult to improve the current situation of higher education through various reforms, which requires joint efforts from various aspects to formulate and gradually implement policies from the national, economic and policy levels.

6 Evaluation and extension of the model

6.1 Advantages of the model

For Model I, the evaluation index selection model based on principal component analysis is used to objectively, reasonably and simply analyze the applicability of each candidate evaluation index, and select the most important evaluation index from several candidate indicators according to the information contribution rate of each index.

For Model II, the health and sustainability index evaluation model of national higher education system based on entropy weight method uses data to illustrate the importance of each index, instead of being dominated by subjective consciousness. Compared with those subjective assignment methods, it is more accurate and more objective, and can better explain the results obtained. It makes the evaluation model of health and sustainability index more convincing and the calculated index can better reflect the health status of the national higher education system.

For Model III, grey prediction requires small sample data, does not require regular distribution of samples, and has a small computational workload, so it has a better prediction effect for the complex system with uncertain factors.

6.2 Disadvantages of the model

For Model I, for in non-gaussian distribution situation, principal component analysis method of principal component may not be optimal. Moreover, if the user has a certain prior knowledge of the observed object and has mastered some characteristics of the data, but cannot intervene in the processing process through parameterization and other methods, the expected effect may not be achieved and the efficiency is not high.

For Model II, ignoring the importance index itself, sometimes to determine index weight and the expected result is far, at the same time, entropy value method does not reduce the dimensions of the evaluation indexes, namely entropy weight method conforms to mathematical laws has strict mathematical sense, but tend to ignore the decision-makers subjective intention;

Entropy weights can be limited if the index changes very little or becomes large or small suddenly.

For Model III, grey prediction based on the predictions of a index rate without considering randomness in the system, medium and long term prediction accuracy is poor.

6.3 Improvements of the model

For Model I, the numerical an average to be processed after processing, can eliminate the influence of partial variable dimension and order of magnitude.

For Model II, subsequent need according to the actual situation, a preference coefficient can be calculated for each, on this basis, the reuse of entropy method for weight calculation, so the importance of the index itself will be collected and the parameter values, the weights of higher accuracy [7].

For Model III, can use other functions on the grey prediction model of the original discrete data column transformation, can improve the smoothness of the discrete data column, broadening the application range of the grey prediction model.

7 Reference

[1] D. E. KNUTH The TEXbook the American Mathematical Society and Addison-Wesley Publishing Qiu Jun, Zhuo Fei, Li Yan. Evaluation model of urban and rural water supply integration based on AHP entropy method [J]. Guangdong water conservancy and hydropower, 2020(07):16-21.

[2] Yang Hao Chang, Ge Hui, Zhang Fan. Discussion on the construction of evaluation index system for high quality development of Higher Education [J]. Education guide, ,2020(10):83-90.

[3] Chen Feiyue. Evaluation of talent cultivation quality of Higher Vocational Education Based on factor analysis [J]. Think tank era, 2019(45):108-109.

[4] Sun Jihong, Wang Chuanhui, Zhao Xiaoyun. Research on comprehensive evaluation of "double first class" higher education quality based on intuitionistic fuzzy model [J]. Journal of Chaohu University, 2019,21(06):133-139.

[5] Qi Yong, Ou Lingyan. Research on the quality evaluation index system of China's higher education development [J]. Education modernization, 2020,7(46):107-113.

[6] Ding Hongwei, Tang Ying, Cao Tiantian. Comparative study on higher education quality evaluation mechanism between China and Japan [J]. Journal of Southwest Jiaotong University (SOCIAL SCIENCE EDITION), 2020,21(03):46-52.

[7] Lin Tongzhi, Tang Guoqiang, Luo Shengfeng, et al. Application of comprehensive evaluation based on improved entropy weighting method and TOPSIS Model [J]. Journal of Guilin Uni