J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 5803--5822

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

The Construction of Educational Management Performance Evaluation System Based on the Hierarchical Analysis Method

Jingjing Wang[1, ✉]

[1] *Organization Department, Zhejiang Guangsha Vocational and Technical University of Construction, Dongyang, Zhejiang, 322100, China*

ABSTRACT
The introduction of performance evaluation in the educational management of colleges and universities is conducive to the formation of result-oriented concepts and management methods of student educational management. In this paper, we select the indicators of educational management conditions, processes and results to design the performance evaluation index system of educational management. Using the hierarchical analysis method, the eigenvectors and maximum eigenvalues are calculated to determine the weights of each index element of the index system. Then apply the gray correlation method to evaluate the educational management performance of the five universities by calculating, one by one, the absolute difference between each indicator sequence (comparative sequence) and the corresponding element of the reference sequence of the object to be evaluated after the data are dimensionless. The analysis found that, according to the formula for calculating the degree of correlation between the actual level of educational management performance and the ideal educational management performance situation, the comprehensive correlation degree of each sample of colleges and universities in the five stages is Z = (0.3333, 0.3951, 0.4600, 0.5031, 0.5946, 1.0000), and the rankings of colleges and universities in terms of the performance of educational management from the highest to the lowest are Academy 4, Academy 2, Academy 5, Academy 3, Academy 1. HEI 3 and HEI 1 should reflect on the shortcomings, enhance the digital construction of teaching informationization, deepen the collaboration between schools and enterprises, and improve the performance of educational management of colleges and universities.

Keywords: hierarchical analysis, gray correlation analysis, dimensionless, educational management performance

1. Introduction

With the continuous development of society and the globalization of education, educational management is advancing with the times. Educational management refers to the effective management and operation of educational institutions through management functions such as planning, organizing, coordinating, monitoring and evaluating [1]. Educational management includes

✉ Corresponding author.
E-mail address: wjjzugc@163.com (J. Wang).
Received 15 January 2024; Revised 20 February 2024; Accepted 30 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-322

management at all levels, from the campus level to the education system level, covering resource allocation, personnel management, curriculum design, and student management [2-3]. The contents of its management system are numerous, and problems such as backward educational management concepts, insufficient ability of management personnel, and imperfect management mode have appeared [4]. In order to strengthen the campus management, increase the reform of personnel system and distribution system, and establish a vibrant school management mechanism; at the same time, stimulate the spirit of love for school, love of duty, dedication and devotion of all staff, mobilize the heart and creativity of staff, and guide staff to concentrate on teaching, management, service and environment [5-7]. Therefore, there is a need to evaluate the achievements and effectiveness of the educational management process, i.e. performance evaluation.

Performance evaluation is an important means of comprehensive evaluation and scientific management of educational institutions. It can help educational institutions to plan and develop better and provide scientific guidance and decision support for their development. Through performance evaluation, we can find out the problems and shortcomings of education management, and targeted to improve and progress, so as to improve the level of education management and optimize the process of education management. At the same time, through performance evaluation, the duties and responsibilities of educational management personnel are clarified to promote the better performance of their duties, thus forming a better teaching environment [8-10]. In addition, in the process of educational management performance evaluation, it can also promote the in-depth development of educational and teaching reform. The main contents of the evaluation are (1) educational management ability, such as the professional knowledge, organization and coordination ability, communication ability of educational management personnel, including but not limited to the understanding and mastery of educational regulations and policies, and the ability to organize teachers' training and meetings [11-13]. (2) The effectiveness of educational management, such as the educational administrators in the development of teaching plans, curriculum management, teacher evaluation and student management, including but not limited to the implementation of teaching plans, teacher evaluation of teaching effectiveness and student development [14-15]. (3) Educational management performance, such as the performance of educational administrators in educational quality control, teaching resource allocation, and campus safety management, including but not limited to the results of educational quality assessment, the maintenance of classroom facilities, and the implementation of campus safety precautions [16-18]. The evaluation system of educational management performance is huge, and in the development of information technology today, it is necessary to improve the efficiency and accuracy of evaluation.

This paper adopts the hierarchical analysis method to construct the index system for evaluating the educational performance of colleges and universities, and combines the information to select the condition level, process level, and outcome level indicators. Then according to the established indicator system, design the expert consultation form, invite experts to fill in the consultation form and construct the judgment matrix, calculate the eigenvectors and maximum eigenvalues and carry out the consistency test, and finally determine the indicator weights. Then the grey correlation analysis is introduced to quantify the educational performance management of the five universities, and the indicator data are dimensionless. The correlation coefficient is calculated to get the value of the comprehensive correlation degree between the actual development status of educational management performance and the ideal educational performance management status.

2. Construction of AHP-based performance evaluation system for education management

2.1. Theory of Hierarchical Analysis (AHP)

AHP method is a hierarchical empowerment decision analysis method, which can consider the combination of quantitative and qualitative at the same time, and at the same time has the advantages of flexibility and simplicity, which has been widely used in many disciplines, such as natural sciences, social sciences and engineering technology. The method first decomposes a complex problem with multiple constituent factors, then decomposes the elements again according to the primary and secondary relationships, and arranges the results of the re-decomposition according to different levels to obtain a multi-objective, multi-level model, and then form a hierarchical hierarchical structure [19].

2.2.1. Characteristics of the AHP method. The AHP method is based on the nature of complex decision-making problems, influencing factors and the intrinsic connection between them, and uses a simple and flexible calculation method to carry out qualitative and quantitative analyses in order to achieve the solution of complex decision-making problems with multiple objectives, multiple criteria and difficult to quantify. The basic evaluation process is as follows: firstly, according to the top-down order, the complex decision-making problem is divided into the overall objective, sub-objectives and alternatives at each level, and is processed according to the hierarchical relationship of the layers, corresponding to the layer, criterion layer and program layer, respectively. Secondly, the relationship between the indicative elements of each level and the indicative elements of the previous level, that is, the priority, is sought; finally, the weighted sum method is used to integrate the priority of each level hierarchically and obtain the weight of each option on the decision-making objective. On this basis, a program with higher weight is proposed.

AHP method is a system analysis method that organically combines qualitative and quantitative. For the multi-objective, multi-indicator, difficult to quantify multi-objective, multi-indicator, difficult to quantify difficulties, based on the expert's cognition, it transforms them into single-objective, easy-to-quantify multi-objective, multi-indicator, difficult-to-quantify difficulties and proposes the best solution. The method does not require much quantitative data and is characterized by simplicity, convenience and practicability. However, the AHP method mainly relies too much on the subjective judgment of experts and is mostly a qualitative analysis, resulting in low credibility. In a system with a large number of indicator factors, the huge amount of calculation will not only bring great difficulty to decision-making, but also change the accuracy of decision-making greatly.

2.2.2. Calculation steps of the AHP method. The analysis process of the AHP method is shown in Figure 1:

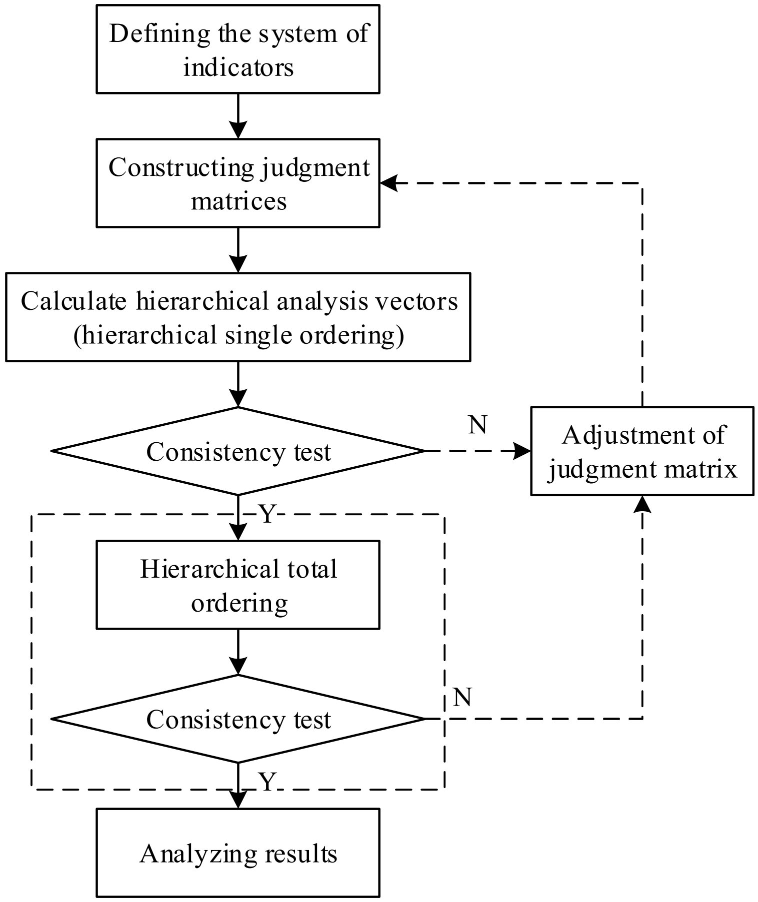

Fig. 1. AHP analysis process

The calculation steps of AHP method are mainly divided into the following four steps:
1) Construct the hierarchical structural model. The standard AHP structural model is mainly divided into three levels, as shown in Figure 2. It can be understood from the figure:

(1) Goal layer: it is the first level of the constructed structural model, there is only one factor, and it is the preset goal in the comprehensive evaluation.

(2) Guideline layer: the guideline layer is in the middle of the structural model, containing the intermediate links involved in achieving the goal, which is a layer constructed by decomposing the influencing factors according to the goal layer, and can be composed of multiple sub-indicators. The criterion layer contains all the attributes of the alternatives, that is, the assessment indicators of the alternatives.

(3) Program Layer: The Program Layer is at the bottom of the structural model and contains the full range of alternative programs, measures, etc. to achieve the final objective.

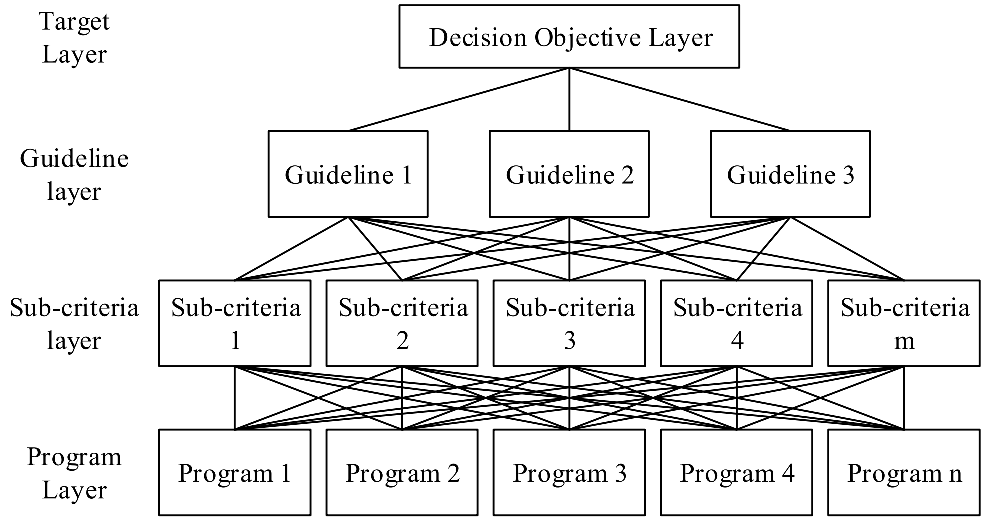

Fig. 2. AHP structure

2) Construction of judgment matrix. AHP needs to construct a judgment matrix, using the way by comparing two indicator factors with each other. Hierarchical analysis structural model uses the upper and lower levels of units corresponding to each other, the upper level of unit factors as constraints, through the same level of factors of two to two comparisons, so as to derive the results, the formation of judgment matrix. The key link is the combination of qualitative and quantitative analysis.

That is, judgment matrix U is:

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ u_{21} & u_{22} & \cdots & u_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nm} \end{pmatrix} \tag{1}$$

Where U represents the target factor of the previous level of the first indicator level factor $U_1, U_2, \cdots, U_n$, which is the decision goal of the first indicator level, and u_{ij} represents the importance of the first indicator level factors U_i and U_j relative to the decision goal U.

The influencing factors in the same level are compared with each other, and the importance of the former factor relative to the latter is described in qualitative terms, and the importance is scaled by the numbers 1 to 9. This method was first used in the allocation of resources, coal mine safety work, etc., and now it can also be used in the safety evaluation of existing buildings, selection and optimization work.

In order to ensure that the resulting judgment matrix can pass the consistency test, it is necessary to use a unified metric to assess the relative importance of each factor, and the 1-9 scale is the most commonly used method to measure the relative importance of each factor:

3) Hierarchical single ranking and consistency test. Hierarchical single sorting is to determine the weights of the index elements to some extent. The method first finds the largest characteristic root of the judgment matrix λ_{max} and its corresponding normalized eigenvector W_i, and then takes it as the weight vector of the corresponding evaluation unit $W = (W_1, W_2, \ldots, W_n)^T$.

The geometric square root of each row vector of the matrix U is then normalized, and the weights and eigenvectors W of each evaluation index can be obtained through equation (2):

$$W_i = \frac{\left(\prod_{j=1}^{n} u_{ij}\right)^{\frac{1}{n}}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} u_{ij}\right)^{\frac{1}{n}}} \tag{2}$$

Normalize the feature vectors to get the weight vector:

$$W = (W_1, W_2, \ldots, W_n)^T \tag{3}$$

The largest characteristic root of the judgment matrix:

$$\lambda_{\max} = \sum \frac{(UW)_i}{nW_i} \tag{4}$$

The judgment matrix U built using the above method may not be a consistency matrix in general, so a consistency test is needed to check whether the weight vector obtained by this method is valid. The following are the specific steps of consistency test:

(1) Calculate the consistency index CI:

$$CI = \frac{\lambda_{\max} - n}{n-1} \tag{5}$$

Where the largest characteristic root of judgment matrix U is $\lambda_{\max}$ and the order of judgment matrix U is n.

(2) Determine the random consistency index RI.

(3) Calculate the consistency ratio CR:

$$CR = \frac{CI}{RI} \tag{6}$$

When $CR \le 0.1$, the consistency test passes; when $CR > 0.1$, the consistency test fails and the judgment matrix U needs to be corrected.

4) Hierarchical total ranking and consistency test. Assuming that the first level indicator layer U contains m indicator factors, respectively $U_1, U_2, \cdots, U_m$, and the weight of each indicator factor corresponding to the target layer is $u_1, u_2, \cdots, u_m$, and the second level indicator layer A includes n indicator factors, respectively $A_1, A_2, \cdots, A_n$, and the weight of one indicator factor $U_k (k = 1, 2, \cdots, m)$ corresponding to the first level indicator layer U is $a_{ik}, a_{2k}, \cdots, a_{nk}$, that is, we can get the total ranking weight of the target layer corresponding to the secondary indicator layer A as $\sum_{k=1}^{m} u_k a_{1k}$, $\sum_{k=1}^{m} u_k a_{2k}$, ..., $\sum_{k=1}^{m} u_k a_{nk}$. Therefore, the percentage of consistency in the total ranking of sub-indicator stratum A is:

$$CR = \frac{\sum_{k=1}^{n} u_k CI_k}{\sum_{k=1}^{n} u_k RI_k} = \frac{u_1 CI_1 + u_2 CI_2 + \ldots + u_k CI_k}{u_1 RI_1 + u_2 RI_2 + \ldots + u_k RI_k} \tag{7}$$

When $CR \le 0.1$, i.e., the hierarchical single ranking of the first level of indicator level U, i.e., the result of the hierarchical total ranking passes the consistency test, the results obtained are reliable.

2.2. Indicator System for Evaluating Educational Management Performance

Performance budgeting is a goal-oriented, project-cost-measured, performance-assessment-centered budgetary system, specifically, a budgetary system that closely integrates increases in resource allocation with improvements in performance. Accordingly, performance evaluation is to improve the effectiveness and efficiency of public utilities through assessment so that government departments can do things rationally and efficiently. The introduction of performance evaluation in the management of student education in local colleges and universities is conducive to strengthening the relationship between objectives and results and the effectiveness of their results in the management of public expenditure on student education, and to forming a result-oriented management concept and management style in order to improve the efficiency, effectiveness and level of management of local student education. The functions and duties of colleges and universities to serve the local economic construction are more prominent and obvious, and only by introducing and constantly strengthening the concept of performance evaluation, the student education of colleges and universities can more reflect their regional characteristics and directly serve the economic and social development.

On the basis of the connotation of the indicators for the assessment of the quality of students' education in the past, the performance evaluation indicators of students' education in local universities are constructed by using the questionnaire survey method and the hierarchical analysis program method (AHP) with scholars and experts, education administrators, school educators and students as the target respondents. The indicator structure adopts the conceptual model of “input-process-output” of system analysis, and the hierarchical system of indicators is determined based on the relationship of indicator attributes. The indicator system is divided into three levels, with the first-level indicators covering the input, process and output of the education performance responsibility system, involving the background resources, operation history and results of student education, which is the main field of the system construction; the second-level indicators include the dimensions of education resources, teaching staff, student quality, teaching management, etc., which belong to the three fields of input, process and output. Tertiary indicators are subdivided for the connotation of the secondary indicators, and are divided into different detailed indicators as the standard of the evaluation system.

According to the different types of student education in local colleges and universities, this paper constructs the evaluation index system. Student education performance evaluation is mainly composed of four elements, one is the main body of the implementation of the assessment; the second is the assessment object as the object of the assessment; the third is the assessment tool, that is, the evaluation index and so on. For particularly important content should be used in a variety of ways to carry out effective.

Communication to ensure that the communication is in place. The final constructed evaluation indicators are shown in Table 1.

Table 1. The design evaluation index of this paper

<table>
<tr><th colspan="2">Primary indicator</th><th colspan="2">Secondary indicator</th><th colspan="2">Tertiary index</th></tr>
<tr><td rowspan="7">A</td><td rowspan="7">Condition level</td><td rowspan="3">A1</td><td rowspan="3">Education resources</td><td>A11</td><td>Teaching basic facilities</td></tr>
<tr><td>A12</td><td>Funding for teaching</td></tr>
<tr><td>A13</td><td>Education requirements</td></tr>
<tr><td rowspan="4">A2</td><td rowspan="4">Source quality</td><td>A21</td><td>Professional background of examinee</td></tr>
<tr><td>A22</td><td>Working experience requirements</td></tr>
<tr><td>A23</td><td>The subjects are professional</td></tr>
<tr><td>A24</td><td>Performance of the combined test</td></tr>
</table>

				A 25	The age, professional and professional title structure is reasonable
		A 3	Teacher team	A 31	Combination of full-time and part-time work
				A 32	Full-time teachers have the same or similar professional master's degree
				A 33	Part-time teachers have outstanding results or rich practical experience in relevant professional areas
				A 34	Complete specifications of training programs, training plans and syllabus
B	Process level	B 1	Course teaching	B 11	Complete specifications of training programs, training plans and syllabus
				B 12	The course sets reasonable science, embodies the professional characteristics, the practicality, the applicability strong
				B 13	Teaching materials, courseware and experimental procedures to meet the teaching requirements
				B 14	Academic leading courses or lectures
				B 15	Teaching conditions are good, and there is a teaching method for master's degree
		B 2	Teaching management	B 21	The management institutions are sound and there is a special person responsible for teaching
				B 22	Implement the training plan and teaching plan, the examination is strict
				B 23	Students have been studying for more than half a year
				B 24	Adopt the classification system, and cultivate according to the characteristics of the students.
C	Results level	C 1	Teaching effect	C1 1	The test scores are reasonable
				C1 2	Proficient in professional knowledge
		C 2	Dissertation	C2 1	The topic is closely combined with professional development actual demand
				C2 2	The research content is realistic and operational
				C2 3	Ability to fully reflect the problem of students
		C 3	Social evaluation	C3 1	Student satisfaction
				C3 2	The employer's ability to evaluate students' business ability
				C3 3	Professional rewards and achievements within two years after graduation

2.3. Calculation of indicator weights

2.3.1. Hierarchical analysis to solve the weighting process. With regard to the established performance indicator system, the process of determining the weights of each indicator element of the indicator system using the hierarchical analysis method is specifically analyzed.

1) According to the establishment of the indicator system, the design of expert consultation table

The same level of indicators on the importance of the previous level of the two-two comparison, the scale of measurement is divided into levels, respectively, absolutely important, very important, more important, slightly important, equally important, corresponding to the value of 9, 7, 5, 3, 1, respectively.

Each indicator is described in the expert consultation form to allow experts to understand the meaning of the indicator more comprehensively and obtain more objective values.

2) Invite experts to fill out the consultation form and construct a judgment matrix

The determination of the consulting object (experts) is a key issue, which is directly related to the quality and credibility of the evaluation work. In this consulting survey, teachers are the main focus. Generally speaking, it is appropriate to consult experts with 5-10 people, and 7 consultation forms were distributed to experts in this survey. After consulting the experts to get the data, the corresponding judgment matrix is constructed.

3) Calculate eigenvectors and maximum eigenvalues and perform consistency test

Calculate the eigenvectors and maximum eigenvalues of the matrix obtained above, and do consistency testing of each expert's opinion with random indicators, consistency indicators, consistency ratio and consistency test for each level of order. The consistency ratio $CR < 0.1$ is usually considered reasonable. If $CR \geqslant 0.1$, necessary adjustments are made to reduce the error after communicating with each expert. Finally, the opinions of the expert group are synthesized. The method of synthesis is generally the arithmetic average or geometric average of the weights of the same indicator given by each expert, and in this paper, the arithmetic average is used to synthesize the weights obtained by each expert.

2.3.2. Calculation and determination of indicator weights. Determine the weights of the first-level indicators, taking expert 1 as an example, first construct a judgment matrix based on the values of the first-level indicators in the expert 1 questionnaire as shown in Table 2.

Table 2. The judgment matrix of the weight of a primary indicator

	A	B	C
A	1	5	4
B	0.2	1	0.25
C	0.25	0.25	1

Using the software to carry out calculations to get the maximum eigenvalue of the judgment matrix of the expert 1 on the first-level indicators of the weight of each indicator.

Then conduct consistency test on the judgment matrix, and use the consistency test formula to get the consistency index CI and consistency ratio CR.

Through the above steps, the weights of the indicators, the maximum eigenvalue max and the consistency ratio CR of the judgment matrix of the other six experts are obtained. it should be noted that, when the consistency ratio $CR \geqslant 1$ of individual experts occurs during the specific calculation process, the judgment matrix is adjusted by a small margin after communicating with the experts to pass the consistency test, so as to reduce the error.

Using the software to calculate, the weights and consistency ratios CR given by each expert for the first-level indicators are shown in Table 3.

Table 3. The weight and consistency ratio of the index is given by experts

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	Arithmetic mean
WA	0.3456	0.3264	0.3464	0.2437	0.2619	0.3687	0.3461	0.3198
WB	0.3511	0.3106	0.3019	0.3268	0.3105	0.3831	0.3043	0.3269
WC	0.3465	0.3324	0.3604	0.3505	0.3462	0.3752	0.3618	0.3533
CR	0.0355	0.0344	0.036	0.0373	0.0358	0.077	0.0368	

λmax	3.0949	3.0967	3.0984	3.098	3.0948	3.1377	3.0977	

It can be seen that the judgment matrix CR of each expert for the first-level indicators is less than 0.1, through the consistency test, and then through the arithmetic average to get the comprehensive weight coefficients of the experts for the first-level indicators, and get the weights of the first-level indicators A (condition level), B (process level), and C (outcome level) on the total objectives as 0.3198, 0.3269, and 0.3533.

Determine the weight coefficients of the second-level indicators, according to the same method described above, use the software to calculate to get the weights and consistency ratio CR of A1 (educational resources), A2 (quality of student source), A3 (teacher team) given by each expert and the maximal eigenvalue max of their judgment matrices, as shown in Table 4.

It can be obtained that the weights of indicators A1 (educational resources), A2 (quality of student population), and A3 (teacher workforce) are 0.3483, 0.3273, and 0.3244, respectively.

Table 4. Weight and consistency ratio CR

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	Arithmetic mean
WA1	0.322	0.398	0.284	0.349	0.386	0.361	0.338	0.3483
WA2	0.308	0.328	0.359	0.302	0.334	0.327	0.333	0.3273
WA3	0.281	0.331	0.367	0.335	0.333	0.336	0.288	0.3244
CR	0	0	0	0	0	0	0	
λmax	3	3	3	3	3	3	3	

Similarly, the weights and consistency ratios CR and the maximum eigenvalue max of their judgment matrices given by each expert for B1 (course teaching), B2 (teaching management), are obtained by using the software, as shown in Table 5. The single sort weights of B1 (course teaching) and B2 (instructional management) can be obtained as 0.5103 and 0.4897.

Table 5. Weight and consistency ratio CR

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	Arithmetic mean
WB1	0.527	0.483	0.512	0.513	0.544	0.543	0.45	0.5103
WB2	0.459	0.454	0.462	0.522	0.487	0.531	0.513	0.4897
CR	0.022	0.023	0.026	0.03	0.025	0.024	0.023	
λmax	2.035	2.105	2.007	2.086	2.047	2.08	2.101	

The weights and consistency ratios CR of C1 (teaching effectiveness), C2 (dissertation), C3 (social evaluation) given by each expert and the maximum eigenvalues of their judgment matrices are obtained by using the software, as shown in Table 6.

The single-ranking weights of C1 (teaching effectiveness), C2 (dissertation), and C3 (social evaluation) can be obtained as 0.3053, 0.3833, and 0.3114.

Table 6. Weight and consistency ratio CR

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	Arithmetic mean
WC1	0.296	0.314	0.326	0.296	0.33	0.284	0.291	0.3053
WC2	0.406	0.419	0.306	0.375	0.383	0.391	0.403	0.3833
WC3	0.316	0.324	0.321	0.283	0.315	0.323	0.298	0.3114

CR	0.024	0.027	0.025	0.024	0.021	0.021	0.029	
λmax	3.074	3.056	3.181	3.009	3.012	3.036	3.027	

Finally, the total ranking weights of the secondary indicators to the overall goal can be obtained, as shown in Table 7.

Table 7. The secondary index is the total order weight coefficient of the total object

Primary indicator	Global weight	Secondary indicator	Global weight
A	0.3198	A1	0.1114
		A2	0.1047
		A3	0.1037
B	0.3269	B1	0.1668
		B2	0.1601
C	0.3533	C1	0.1079
		C2	0.1354
		C3	0.11

The weight coefficients of the three levels of indicators were calculated according to the same method described above. The final evaluation indicator system was obtained as shown in Table 8. Among the three-level indicators, the highest global weight is C12, with a weight of 0.0587.

Table 8. Management performance evaluation system weight

Primary indicator	Global weight	Secondary indicator	Global weight	Indicator	Global weight
A	0.3198	A1	0.1114	A11	0.0328
				A12	0.0337
				A13	0.0449
		A2	0.1047	A21	0.0197
				A22	0.0189
				A23	0.0244
				A24	0.0203
				A25	0.0214
		A3	0.1037	A31	0.0254
				A32	0.0239
				A33	0.0265
				A34	0.0279
B	0.3269	B1	0.1668	B11	0.0316
				B12	0.0313
				B13	0.0387
				B14	0.0299
				B15	0.0353

		B2	0.1601	B21	0.0416
				B22	0.0407
				B23	0.0373
				B24	0.0405
C	0.3533	C1	0.1079	C11	0.0492
				C12	0.0587
		C2	0.1354	C21	0.0463
				C22	0.0439
				C23	0.0452
		C3	0.11	C31	0.0349
				C32	0.0396
				C33	0.0355

3. Empirical analysis of the application of evaluation systems

3.1. Data sources

According to the establishment of the indicator system to create a questionnaire, the survey object for college managers, respectively, five college samples issued questionnaires, retrieved questionnaires to remove unreasonable data. In order to improve the accuracy of the results, this paper uses the average of the sample survey data at all stages, then the 29 indicators in the five stages of data shown in Table 9.

Table 9. The Statistical data at all stages

Index factor	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Ideal value
A11	2.0652	2.6087	2.9103	3.2317	3.9872	5
A12	1.5219	2.3777	3.1392	3.3281	4.4113	5
A13	1.5495	1.9534	3.1381	3.8671	4.0403	5
A21	1.8278	2.2175	2.9814	3.5863	4.517	5
A22	1.5257	2.6873	3.1399	3.2694	4.0064	5
A23	1.9036	2.6228	3.0737	3.4665	4.5995	5
A24	2.0934	2.0098	3.3574	3.6152	4.3704	5
A25	1.8556	2.3002	3.3225	3.238	4.3145	5
A31	2.1943	2.1188	3.1923	3.2978	4.4358	5
A32	1.5943	2.3792	3.0214	3.7566	4.577	5
A33	1.6315	2.2193	3.264	3.6082	3.9211	5
A34	1.8565	2.0375	2.9757	3.7248	4.5658	5
B11	1.9621	2.3414	3.2669	3.2469	4.4885	5
B12	1.6842	2.4164	3.3343	3.6319	3.9466	5
B13	1.9239	2.5697	3.5851	3.2728	4.592	5
B14	1.9774	2.5071	3.1496	3.1446	4.5851	5
B15	1.9977	2.3207	3.2728	3.513	3.821	5
B21	1.5461	2.2672	3.1355	3.3126	4.5629	5
B22	2.1188	2.1974	2.9096	3.3728	4.1317	5
B23	1.8185	2.3334	3.2613	3.6492	4.4318	5
B24	2.1846	2.1482	3.468	3.1497	3.8786	5
C11	1.8253	2.1352	2.9561	3.6162	4.4577	5
C12	1.8764	2.399	2.8395	3.1747	4.2174	5
C21	1.5079	2.184	3.5079	3.6949	4.409	5
C22	2.2036	2.0789	2.913	3.3811	4.1178	5
C23	2.166	2.0523	3.4979	3.4933	4.4915	5
C31	1.8369	2.6708	3.2164	3.2489	4.3977	5
C32	1.6329	2.2255	3.4389	3.7572	4.411	5
C33	2.0245	2.0511	3.0267	3.3065	4.2703	5

According to the algorithm for 29 indicators in 5 stages of the original data and the ideal value of each indicator is dimensionless, then the standardized results are shown in Table 10.

Table 10. The dimensionless quantification processing data

Index factor	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Ideal value
A11	0	0.2065	0.4321	0.5152	0.7176	1
A12	0	0.1346	0.3994	0.5754	0.7259	1
A13	0	0.1543	0.3803	0.5168	0.5993	1
A21	0	0.1691	0.4582	0.5059	0.5854	1
A22	0	0.2447	0.4225	0.4876	0.6545	1
A23	0	0.2238	0.4561	0.5708	0.6425	1
A24	0	0.1208	0.4484	0.4793	0.5953	1
A25	0	0.1924	0.4645	0.558	0.6513	1
A31	0	0.221	0.4051	0.528	0.6189	1
A32	0	0.2315	0.4003	0.5533	0.7201	1
A33	0	0.2147	0.4323	0.4683	0.6646	1
A34	0	0.1433	0.3868	0.5313	0.7243	1
B11	0	0.1404	0.4145	0.4894	0.6554	1
B12	0	0.171	0.3845	0.5279	0.6355	1
B13	0	0.1779	0.4204	0.502	0.6673	1
B14	0	0.1658	0.3827	0.5448	0.662	1
B15	0	0.2278	0.3743	0.4639	0.6202	1
B21	0	0.1489	0.4302	0.4928	0.6598	1
B22	0	0.2522	0.4371	0.519	0.6392	1
B23	0	0.1449	0.4444	0.4747	0.6465	1
B24	0	0.1579	0.4251	0.5707	0.683	1
C11	0	0.1673	0.4355	0.4966	0.6732	1
C12	0	0.22	0.4356	0.4885	0.7356	1
C21	0	0.1764	0.4728	0.5294	0.6213	1
C22	0	0.2163	0.4631	0.5139	0.732	1
C23	0	0.2424	0.3913	0.5778	0.7315	1
C31	0	0.1805	0.4453	0.5296	0.7005	1
C32	0	0.1984	0.4216	0.5549	0.6051	1
C33	0	0.1708	0.4435	0.5435	0.6383	1

3.2. Gray correlation analysis comprehensive evaluation model

Gray correlation analysis is a factor analysis method, which is characterized by “less data modeling, small samples, poor information uncertainty”, mainly by comparing the geometric relationship of the system data series curve to analyze the degree of association of the variables in the system, gray correlation is used to describe the size and strength of the variables. Gray correlation analysis does not require a large sample size, nor does it require a typical distribution pattern in order to analyze the results, but the results of the quantitative analysis almost coincide with the qualitative analysis. Gray correlation analysis mainly quantifies the degree of geometric similarity between the curves, the closer the curves are, the greater the degree of correlation between the corresponding sequences, on the contrary, the smaller. Gray correlation analysis mainly quantifies the geometric similarity between the curves represented by the sequences and the curves represented by the reference sequences, thus describing the degree of correlation between the two sets of figures through quantitative results. Currently, gray correlation analysis has a wide range of applications in economic, social, mining, agriculture, transportation and other fields.

Gray correlation analysis does not require too much data about the sample or the data of regular samples, and it is suitable for the system whose information is not very clear. Thus, it makes up for the regret of needing large samples and precise distribution law to do system analysis, which is based on the gray correlation degree calculated from the sample data as a model, thus judging the relationship between the factors. Gray correlation analysis is small, suitable for research with uncertain information and small sample size, and the results obtained through gray correlation analysis are basically consistent with the actual.

3.2.1. Gray correlation analysis. Regression analysis, analysis of variance, principal component analysis and other methods are commonly used to analyze the mathematical statistics of the system, but these methods have certain shortcomings:

1) There must be a sufficiently large amount of data, otherwise the data cannot be statistically analyzed.

2) The probability distribution that the sample needs to obey is a typical distribution, and the factors are independent of each other, but they are linear. These factors consist mainly of the relationship between the data and the data characterizing the system, which is usually difficult to satisfy.

3) The amount of data required for the calculation is large and usually depends on the computer [20].

4) The results may be inconsistent with quantitative and qualitative analyses, leading to distortions and reversals of system relationships and laws. In China, the methods of statistical data are very limited, and there are many problems with the existing statistical data, mainly because of the large gray scale of the data, coupled with human reasons, the conclusion of the ups and downs are all phenomena occurring in a lot of data and do not have a typical distribution law. It is usually difficult to solve these problems using traditional mathematical and statistical methods.

Gray correlation analysis method is also applicable to the small sample size and the presence or absence of samples, quantitative results and qualitative analysis of the results of inconsistent data statistics, it has a small amount of calculations, easy to statistics and other characteristics. Therefore, it better compensates the disadvantages brought by using mathematical statistics for system analysis.

3.2.2. Modeling steps. The modeling process of gray correlation analysis is as follows:

1) Set n data series to form the following matrix:

$$\left(x_1', x_2', \cdots, x_n'\right) = \begin{bmatrix} x_1'(1) & x_2'(1) & \cdots & x_n'(1) \\ x_1(2) & x_2(2) & \cdots & x_n(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1'(m) & x_2'(m) & \cdots & x_n'(m) \end{bmatrix} \tag{8}$$

$$x_i' = \left(x_i'(1), x_i'(2), \cdots, x_i'(m)\right)^T \tag{9}$$

where m is the number of indicators, $i = 1, 2, \cdots, n$

2) Determine the reference data column. Determine the reference series should be an ideal comparison standard, can be the optimal value (or the worst value) of each indicator to constitute the

reference series, can also be based on the evaluation purpose to choose other reference values. Mark as:

$$x_0' = \left(x_0'(1), x_0'(2), \cdots, x_0'(n)\right) \tag{10}$$

3) Programmability of Indicator Data. When comparing indicator data, the outline of the data may lead to different results in the system due to various factors such as different physical meanings, which may not be easy to make comparisons, or it may be difficult to get the correct conclusions when making comparisons. When performing grey correlation analysis, the data are generally subjected to non-dimensional dimensioning [21].

Commonly used non-dimensional dimensioning methods include homogenization (11), normalization (12) and (13) transformations.

$$x_i(k) = \frac{x_i(k)}{\frac{1}{m}\sum_{k=1}^{m} x_i'(k)} \tag{11}$$

$$x_i(k) = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{12}$$

$$x_i(k) = \frac{x - \overline{x}}{s} \tag{13}$$

The dimensionless data series form the following matrix:

$$\left(x_0(1), x_0(2), \cdots, x_0(m)\right) = \begin{bmatrix} x_0^{(1)} & x_1(1) & \cdots & x_n(1) \\ x_0(2) & x_1(2) & \cdots & x_n(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_0^{(m)} & x_1(m) & \cdots & x_n(m) \end{bmatrix} \tag{14}$$

4) Calculate, one by one, the absolute difference between each indicator sequence (comparison sequence) of the object to be evaluated and the corresponding element of the reference sequence, i.e., $\left|x_0(k) - x_i(k)\right|$ ($k = 1, \cdots, m$, $i = 1, \cdots, n$, n are the number of evaluated objects).

5) Determine the relationship coefficient, i.e:

$$\min_{=1}^{n} \min_{i=1}^{m} \left|x_0(k) - x_i(k)\right| \; And \; \max_{i=1}^{n} \max_{k=1}^{m} \left|x_0(k) - x_i(k)\right| \tag{15}$$

6) Calculate the correlation coefficient. Calculate the correlation coefficients of each interval comparison sequence and the corresponding elements of the reference sequence, respectively:

$$\xi_i(k) = \frac{\min_i \min_k \left|x_0(k) - x_i(k)\right| + \rho \cdot \max_i \max_k \left|x_0(k) - x_i(k)\right|}{\left|x_0(k) - x_i(k)\right| + \rho \cdot \max_i \max_k \left|x_0(k) - x_i(k)\right|} \tag{16}$$

$k = 1, \cdots, m$, where ρ is the discrimination coefficient, $0 < \rho < 1$. If ρ is smaller, the greater the difference in coefficients, the better the discrimination. Usually ρ is taken as 0.5, and when the best (or worst) value of each indicator is used to form a reference data column for the calculation of the correlation coefficient, an improved and simpler method of calculation can be used:

$$\xi_i(k) = \frac{\min_i \min_k \left|x_0'(k) - x_i'(k)\right| + \rho \cdot \max_i \max_k \left|x_0'(k) - x_i'(k)\right|}{\left|x_0'(k) - x_i'(k)\right| + \rho \cdot \max_i \max_k \left|x_0'(k) - x_i'(k)\right|} \tag{17}$$

7) Calculation of correlation. Calculate the average of the correlation coefficients between each indicator of each assessment object and the corresponding element of the reference sequence (comparison sequence), and in order to reflect the relationship between each assessment object and the reference sequence, and to become a correlation sequence, recorded as:

$$\gamma_{ai} = \frac{1}{m}\sum_{k=1}^{m}\zeta_i(k) \tag{18}$$

3.3. *Evaluation results and analysis*

The reference series selected in this paper is the ideal value of each indicator, and the comparison series is the statistical data of 29 indicators in 5 stages, and the correlation coefficient is calculated according to the algorithm after the dimensionless processing, and in the gray theory, usually the researchers and scholars take = 0.5, then the correlation coefficient of the comparison series and the reference series is shown in Table 11. The correlation coefficient of stage 1 is consistent, all of them are 0.3333. The coefficient of stage 2, 3, 4 and 5 is gradually raised, and the mean value is 0.3951, 0.4600, 0.5031 and 0.5946 respectively.

Table 11. The relevance coefficient table

Index factor	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Ideal value
A11	0.3333	0.3768	0.439	0.5156	0.596	1
A12	0.3333	0.4192	0.4611	0.5106	0.6267	1
A13	0.3333	0.4264	0.481	0.5098	0.6078	1
A21	0.3333	0.3985	0.4395	0.4792	0.6078	1
A22	0.3333	0.404	0.4743	0.4895	0.6211	1
A23	0.3333	0.4166	0.4475	0.524	0.6019	1
A24	0.3333	0.3985	0.4446	0.5256	0.5774	1
A25	0.3333	0.3657	0.4687	0.5192	0.5908	1
A31	0.3333	0.369	0.4421	0.5048	0.5631	1
A32	0.3333	0.3989	0.4578	0.5131	0.5859	1
A33	0.3333	0.3886	0.4884	0.5254	0.5818	1
A34	0.3333	0.3794	0.4324	0.4737	0.6151	1
B11	0.3333	0.3851	0.4838	0.4722	0.6121	1
B12	0.3333	0.394	0.4865	0.4973	0.6056	1
B13	0.3333	0.4032	0.4513	0.5174	0.6385	1
B14	0.3333	0.3817	0.4834	0.4933	0.5635	1
B15	0.3333	0.4123	0.4679	0.4869	0.6209	1
B21	0.3333	0.3829	0.4573	0.501	0.5887	1
B22	0.3333	0.4211	0.4433	0.517	0.5919	1
B23	0.3333	0.4041	0.4479	0.4899	0.5599	1
B24	0.3333	0.4127	0.4357	0.4918	0.6114	1
C11	0.3333	0.4049	0.4577	0.4857	0.6189	1
C12	0.3333	0.3761	0.4526	0.5178	0.5622	1
C21	0.3333	0.4176	0.4846	0.4981	0.5603	1
C22	0.3333	0.3905	0.4762	0.4904	0.5683	1

C23	0.3333	0.3729	0.4549	0.5161	0.6242	1
C31	0.3333	0.3783	0.4577	0.4954	0.5732	1
C32	0.3333	0.3522	0.4667	0.5264	0.6076	1
C33	0.3333	0.428	0.4565	0.5041	0.5616	1

Calculating the target level correlation, which is the degree of correlation between the actual level of educational management performance and the ideal educational management performance situation, according to the formula, the combined correlation of each sample university in the five stages is shown below:

Z = (0.3333, 0.3951, 0.4600, 0.5031, 0.5946, 1.0000)

Using the same principle, the calculation was continued to introduce upwards the degree of correlation between the factors of the criterion layer at each stage and the ideal value, and the results are shown in Table 12.

Table 12. The criteria layer relevance degree table

Index factor	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
A1	0.3333	0.3966	0.4833	0.5104	0.5836
A2	0.3333	0.3757	0.4447	0.4913	0.6287
A3	0.3333	0.371	0.4942	0.5284	0.6564
B1	0.3333	0.4148	0.4974	0.523	0.6501
B2	0.3333	0.3681	0.4603	0.515	0.6616
C1	0.3333	0.3874	0.4927	0.4804	0.6062
C2	0.3333	0.3852	0.4495	0.551	0.6683
C3	0.3333	0.3749	0.4704	0.5598	0.6599

According to the same principle of the above calculations, we calculate separately for the five college samples in the fifth stage, then the value of the combined degree of correlation between the actual development status of educational management performance and the ideal educational performance management status of the five college samples is shown in Table 13.

As can be seen from the table, the gray evaluation scores of the five college samples in the fifth stage are ranked as follows: Academy 4 > Academy 2 > Academy 5 > Academy 3 > Academy 1, which indicates that HEI 4 and HEI 2 have a higher degree of correlation with the educational management performance, and have a better performance of educational management performance. Whereas, Academy 3 and Academy 1 have a lower correlation with the ideal settlement landscape, and their educational management performance needs to be improved compared to the other three academies.

Table 13. The comprehensive relevance degree table

	Academy 1	Academy 2	Academy 3	Academy 4	Academy 5
Integrated correlation	0.5834	0.6219	0.5966	0.6571	0.6017
Sort	5	2	4	1	3

4. Conclusion

This paper combines the hierarchical analysis method to design the evaluation index system of educational management performance, and applies the gray correlation analysis method to empirically analyze the educational management performance of five colleges and universities. The final conclusions are as follows:

1) The weights of the first-level indicators A (condition level), B (process level), and C (outcome level) on the global goal are 0.3198, 0.3269, and 0.3533 obtained by hierarchical analysis.

2) Among the three-level indicators, the highest global weight is C12, with a weight of 0.0587. The lowest is A22, with a weight of 0.0189.

3) Calculated for the five college samples in the fifth stage, college 4 and college 2 have higher correlation with educational management performance and better performance in educational management performance. Whereas, HEI 3 and HEI 1 have lower ideal correlation and educational management performance performance needs to be improved compared to the other 3 HEIs.

In view of the educational management problems exposed by college 3 and college 1, they should actively change the original concept of crude educational management, promote collaborative schooling between school and enterprise, and improve the construction of relevant systems to improve the educational management performance of colleges and universities at an early date.

References

[1] Tojimamatovich, J. V. (2023). Digital Transformation of Educational Management System. Web of Semantic: Universal Journal on Innovative Education, 2(4), 202-206.

[2] Connolly, M., James, C., & Fertig, M. (2019). The difference between educational management and educational leadership and the importance of educational responsibility. Educational Management Administration & Leadership, 47(4), 504-519.

[3] Khusni, W., & Mahmudah, F. N. (2020). THe principal's managerial ability in developing effective schools. International Journal of Educational Management and Innovation, 1(2), 99-108.

[4] Lajčin, D. (2021). Difficult situations in educational management. Acta Educationis Generalis, 11(2), 129-144.

[5] Marbun, D. S., Effendi, S., Lubis, H. Z., & Pratama, I. (2020). Role of education management to expediate supply chain management: a case of Indonesian Higher Educational Institutions. International Journal of Supply Chain Management (IJSCM), 9(1), 89-96.

[6] Hallinger, P. (2018). Surfacing a hidden literature: A systematic review of research on educational leadership and management in Africa. Educational Management Administration & Leadership, 46(3), 362-384.

[7] Díez, F., Villa, A., López, A. L., & Iraurgi, I. (2020). Impact of quality management systems in the performance of educational centers: educational policies and management processes. Heliyon, 6(4).

[8] Nojavan, M., Heidari, A., & Mohammaditabar, D. (2021). A fuzzy service quality based approach for performance evaluation of educational units. Socio-Economic Planning Sciences, 73, 100816.

[9] Da Silva, S. L. C., Vieira, L. C., & da Silva, E. P. (2017). Strategy, performance evaluation and process management in higher education institutions. Independent Journal of Management & Production, 8(2), 362-377.

[10] Liu, S., Xu, X., Grant, L., Strong, J., & Fang, Z. (2017). Professional standards and performance evaluation for principals in China: A policy analysis of the development of principal standards. Educational Management Administration & Leadership, 45(2), 238-259.

[11] Nazari-Shirkouhi, S., Mousakhani, S., Tavakoli, M., Dalvand, M. R., Šaparauskas, J., & Antuchevičienė, J. (2020). Importance-performance analysis based balanced scorecard for performance evaluation in higher education institutions: an integrated fuzzy approach. Journal of Business Economics and Management, 21(3), 647-678.

[12] Ibay, S. B., & Pa-alisbo, M. A. C. (2020). An Assessment of the Managerial Skills and Professional Development Needs of Private Catholic Secondary School Administrators in Bangkok, Thailand. World Journal of Education, 10(1), 149-163.

[13] Altinay, F., Altinay, Z., Dagli, G., & Cifci, M. (2018). Evaluation of administrative and supervisory performance of the directors of Special Education Institutions according to the teachers. Quality & Quantity, 52, 1275-1286.

[14] Englund, H., & Frostenson, M. (2017). Managing performance evaluation uncertainties in schools: When teachers become struggling performers. European educational research journal, 16(6), 885-906.

[15] Datahan, P. J. (2020). School monitoring evaluation and adjustment in public secondary schools: Practices and performance of administrators. Journal of World Englishes and Educational Practices, 2(2), 146-157.

[16] Simpeh, F., & Adisa, S. (2021). Evaluation of on-campus student housing facilities security and safety performance. Facilities, 39(7/8), 470-487.

[17] Sarchahani, Z., Zare, R., Ghanbari, S., & Gholtash, A. (2018). Pathology of Administrators performance evaluation system: The results of Qualitative research. School Administration, 6(2), 264-284.

[18] Alani, F. S., Khan, M. F. R., & Manuel, D. F. (2018). University performance evaluation and strategic mapping using balanced scorecard (BSC): Case study–Sohar University, Oman. International Journal of Educational Management, 32(4), 689-700.

[19] Nguyen Thi The Nguyen,Nguyen Thanh Hoa & Nguyen Thi Lien.(2025).Innovative multi-criteria risk assessment framework for unsanitary, closed landfills: Integrating the fuzzy analytic hierarchy process with a weighted geometric mean approach. Science of the Total Environment178245-178245.

[20] Zhenyi Zhao.(2024).Coordinated Development of Civic Education and Three-Whole Parenting Concepts in Colleges and Universities Based on Gray Correlation Approach.Applied Mathematics and Nonlinear Sciences(1).

[21] Wei Wang,Laitong Cui,Xiaolong Geng,Manhua Luo,Shengchao Yu,Zhenyan Wang... & Hailong Li.(2025).Numerical simulations of tidal beach seawater-groundwater circulations using dimensionless method. Journal of Hydrology132372-132372.