ISSN 2616-5775 Vol. 6, Issue 7: 51-58, DOI: 10.25236/AJCIS.2023.060708

Evaluation of Carbon Emissions Redcution Performance Based on TOPSIS and K-Means Clustering Algorithm

Guanqiao Wang, Heng Zhou, Bohang Lin

School of Electronics and Information, Hangzhou Dianzi University, Hangzhou, Zhejiang, 310018, China

Abstract: *Carbon emissions are silent but deadly poison that threaten the environment human live. To accurately assess each country's level of attitudes and commitment in reducing carbon emissions, the article establishes an evaluation model of carbon emissions reduction performance. First of all, TOPSIS based on the entropy weight method (EWM) is adopted to score 198 countries. Then, K-Means clustering algorithm is used to classify the countries into five distinct categories: Top, Good, Middle, Underachievers, and Poor. The number of countries in each category is 5, 66, 51, 36, and 40 respectively. To verify the rationality of classification, carbon emissions data from 1970 to 2021 around the world is utilized to conduct Spearman correlation analysis.*

Keywords: *Carbon emissions reduction, TOPSIS, K-Means clustering algorithm, Correlation analysis*

1. Introduction

Carbon emissions have root in the rapid advancement of modern civilization. The main cause of these emissions is the burning of fossil fuels such as coal, oil, and natural gas. With demand for energy exploding in recent years, the burning of fossil fuels leads to a surge in carbon emissions. In addition, deforestation also plays a significant role in the increase of carbon emissions. Deforestation releases the carbon trees store into the atmosphere, resulting in higher carbon emissions.

Carbon emissions play the major role in climate change. Excessive carbon emissions have caused serious damage to ecosystem, economy, and society. With temperature rising, the risk of extreme weather events such as drought, wildfires, and floods increases. These disasters have the power to wipe out entire society, destroy infrastructure, and lead to widespread famine and disease[1-3].

Reducing carbon emissions is a necessary condition for human beings to survive on the earth. In recent years, countries around the world have been actively developing new energy sources such as hydrogen energy, solar energy, offshore wind energy to achieve carbon neutrality. However, due to the differences among countries in terms of carbon emissions reduction targets, implementation progress of current carbon emissions reduction policies, existence of carbon emissions reduction plans, etc., it is necessary to build a comprehensive evaluation model to take these differences into account.

He et al. used the Logistic Mean Divided Index (LMDI) to decompose and quantify the contributions of nine factors that affect China's CEEI (account for more than 40% of China's total carbon emissions) increase. Then they used K-Means clustering to group provinces and puts forward targeted recommendations[2]. Feng et al. studied the current situation and influencing factors of carbon emissions from rural buildings in a typical village located in southern China, adopting the emission factor method and LMDI-LEAP model[4].

The article provides a comprehensive analysis of carbon neutral targets and relevant information of different countries around the world. Then TOPSIS based on the EWM is adopted to obtain scores of countries with different carbon emissions reduction commitment and attitudes. On this basis, K-Means clustering algorithm is applied to group the countries into five categories based on their scores: Top, Good, Middle, Underachievers, and Poor. This categorization provides a more intuitive understanding of the attitudes and progress of different countries in reducing carbon emissions.

To demonstrate the rationality of the classification, Spearman correlation analysis is conducted. Five countries are randomly selected from each category and their correlation coefficients are compared to determine the rationality of the classification.

ISSN 2616-5775 Vol. 6, Issue 7: 51-58, DOI: 10.25236/AJCIS.2023.060708

2. Evaluation of Carbon Emissions Reduction Performance

2.1 Selection and quantification of evaluation indicators

From https://zerotracker.net/, the carbon neutral targets and relevant information of various countries in the world are obtained. After integration and analysis, four indicators are systematically selected.

The first indicator is “end_target”. The goal of carbon emissions reduction can reflect its status and confidence in carbon neutrality. The quantification and processing methods of indicators are determined as follows.

To quantify the index, targets are rank from best to worst and noted as 7, 6, 5, 4, 3, 2, 1, 0. The detail is shown in Table 1.

Table 1: Quantization of “end_target”

end_target	Quantized value
Zero carbon	7
Net zero	6
Emissions intensity target	5
Emissions reduction target	4
Reduction v. BAU	3
1.5°C target	2
Other	1
No target	0

The second index is “end_target_year”. The earlier “end_target_year is”, the better carbon emissions reduction measures can be reflected to some extent. This article sets the “end_target_year” index in the positive direction.

The third metric, “end_target_status”, is a crucial indicator as it reflects efforts to reduce carbon emissions of a country in past years. This article ranks them from best to worst and records them as 5, 4, 3, 2 and 1, as shown in Table 2.

Table 2: Quantization of “end_target_status”

end_target_status	Quantized value
Achieved (self-declared)	5
In law	4
In policy document	3
Declaration / pledge	2
Proposed / in discussion	1
None	0

The fourth index is “has_plan”, the quantization is shown in Table 3.

Table 3: Quantization of “has_plan”

has_plan	Quantized value
Yes	1
No	0

The evaluation system of carbon emissions reduction is shown in Figure 1.

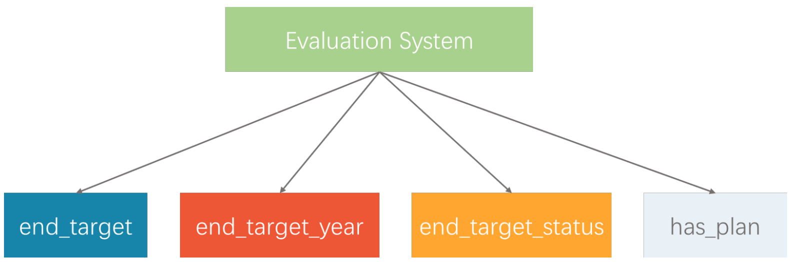


Figure 1: Evaluation system of carbon emissions reduction

ISSN 2616-5775 Vol. 6, Issue 7: 51-58, DOI: 10.25236/AJCIS.2023.060708

2.2 Evaluation of performance based on TOPSIS and the EWM

After quantifying the indicators, the article conducts TOPSIS on 198 countries. To reduce the influence of subjective factors, the EWM is used.

The EWM is one of the weighting methods that measures the dispersion level of different information sources in decision making. It determines the weight of each criterion by calculating the degree of their influence on the decision-making result[5].

Suppose given i index $X_1, X_2, X_3, \ldots, X_i$. and $X_s = \{X_1, X_2, \ldots, X_i\}$. The normalized results are $X_1', X_2', X_3', \ldots, X_i'$, as shown below.

$$X_i' = \frac{X_i - \min(X_s)}{\max(X_s) - \min(X_s)} \tag{1}$$

The formula of information entropy is calculated as follows.

$$e_j = -\ln(n)^{-1} \sum_{i=1}^{n} p_{ij} \ln p_{ij} \tag{2}$$

P_{ij} represents the index ratio of the jth scheme in the ith index.

By normalizing the information utility value, the entropy of each index is as follows.

$$w_j = \frac{1 - E_j}{k - \Sigma E_j} (\mathrm{j} = 1, 2, \ldots, m) \tag{3}$$

The original data matrix is set in the positive direction. Data is preprocessed with regularization and normalization, and the EWM is used to assign weights to the preprocessed indicators. The weight calculation results are shown in Table 4.

Table 4: Results of weight calculation

Index	e	d	weight (%)
end_target_ver	0.983	0.017	10.662
end_target_year_ver	0.981	0.019	11.618
end_target_status_ver	0.966	0.034	21.309
has_plan_ver	0.91	0.09	56.411

TOPSIS is adopted to comprehensively score 198 countries. The parameters follow principles as follows[6].

Calculate the Euclidean distance between the index selected or constructed. The optimal vector is as below.

$$D^+ = \sqrt{\sum_{j=1}^{n} (Z_{ij} - Z_j^+)^2} \tag{4}$$

The worst vector is as follows.

$$D^- = \sqrt{\sum_{j=1}^{n} (Z_y - Z_j^-)^2} \tag{5}$$

The relative proximity to the optimal value is denoted as the comprehensive score of a location's light pollution risk level.

$$W_i = \frac{D_i^-}{D_i^+ + D_i^-} \tag{6}$$

Through TOPSIS, the article obtains the comprehensive scores of 198 countries and then ranks them.

ISSN 2616-5775 Vol. 6, Issue 7: 51-58, DOI: 10.25236/AJCIS.2023.060708

2.3 Classification based on K-Means clustering algorithm

K-Means clustering algorithm is adopted to classify 198 countries[2,7].

K-Means clustering algorithm is used to cluster the results. The algorithm is an unsupervised learning one used to divide a data set into K different clusters or groups. It iteratively divides the sample points in the data set into K clusters, making the distance between each sample point and the center point of the corresponding cluster it the smallest.

According to the clustering coefficient from Spss, the elbow figure is drawn in Figure 2.

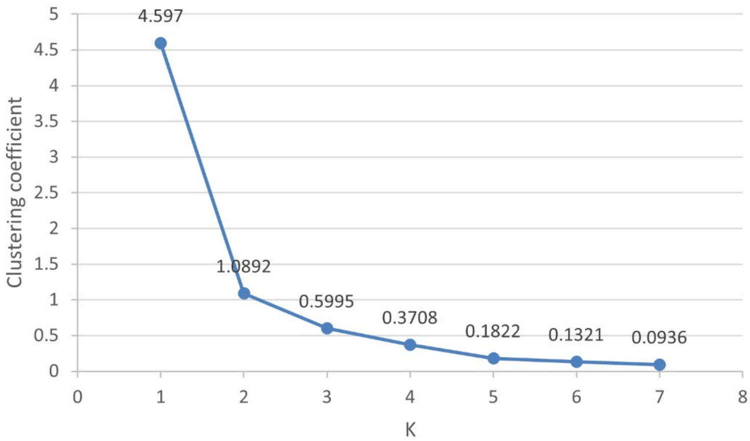


Figure 2: Line diagram of polymerization coefficient (County)

It can be seen from Figure 2 that when the value of K is from 1 to 5, the degree of distortion changes the most. With K exceeding 5, the degree of distortion is significantly reduced, so k is set as 5[8]. Different kinds are as follows: Top, Good, Middle, Underachievers, and Poor.

According to SPSS, 5 countries are included in category of Top: Benin, Madagascar, Gabon, Suriname, and Georgia. Category of Good has 66 countries, represented by: Japan, France, Denmark, Tunisia, etc. 51 countries exist in category of Middle, represented by Monaco, China, Thailand, etc. 36 countries are included in category of Underachievers, represented by Jordan, South Sudan, Fiji, Comoros, etc. Category of Poor has 40 members, such as Timor-Leste, Vanuatu.

The result is shown in Figure 3.

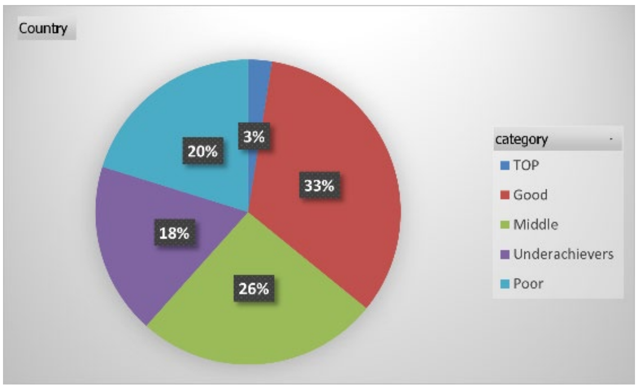


Figure 3: Cluster results

According to the cluster results, most countries, especially developing countries, have relatively low overall evaluation scores due to technical and economic constraints. The purpose of evaluation and classification is not to judge or praise countries, but to guide developing nations in adopting correct attitudes and measures to reduce carbon emissions. However, do not blindly emulate other countries. To promote carbon neutrality in the country, relevant measures should be taken in accordance with local conditions.

ISSN 2616-5775 Vol. 6, Issue 7: 51-58, DOI: 10.25236/AJCIS.2023.060708

3. Correlation Analysis Based on Spearman correlation coefficient

3.1 Correlation between data and selection of correlation coefficient

To demonstrate the rationality of the classification above, five countries from the same category are selected as samples: Angola, Greece, Barbados, Tunisia, and Maldives.

Data of carbon emissions from 1970 to 2021 around the world is collected.

The change curves of data (the total carbon dioxide data of each country, the fossil carbon dioxide data per GDP of each country, and the fossil carbon dioxide data per capita of each country) are drawn as shown in Figure 4 (In order to facilitate data comparison, three data have been enlarged before drawing).

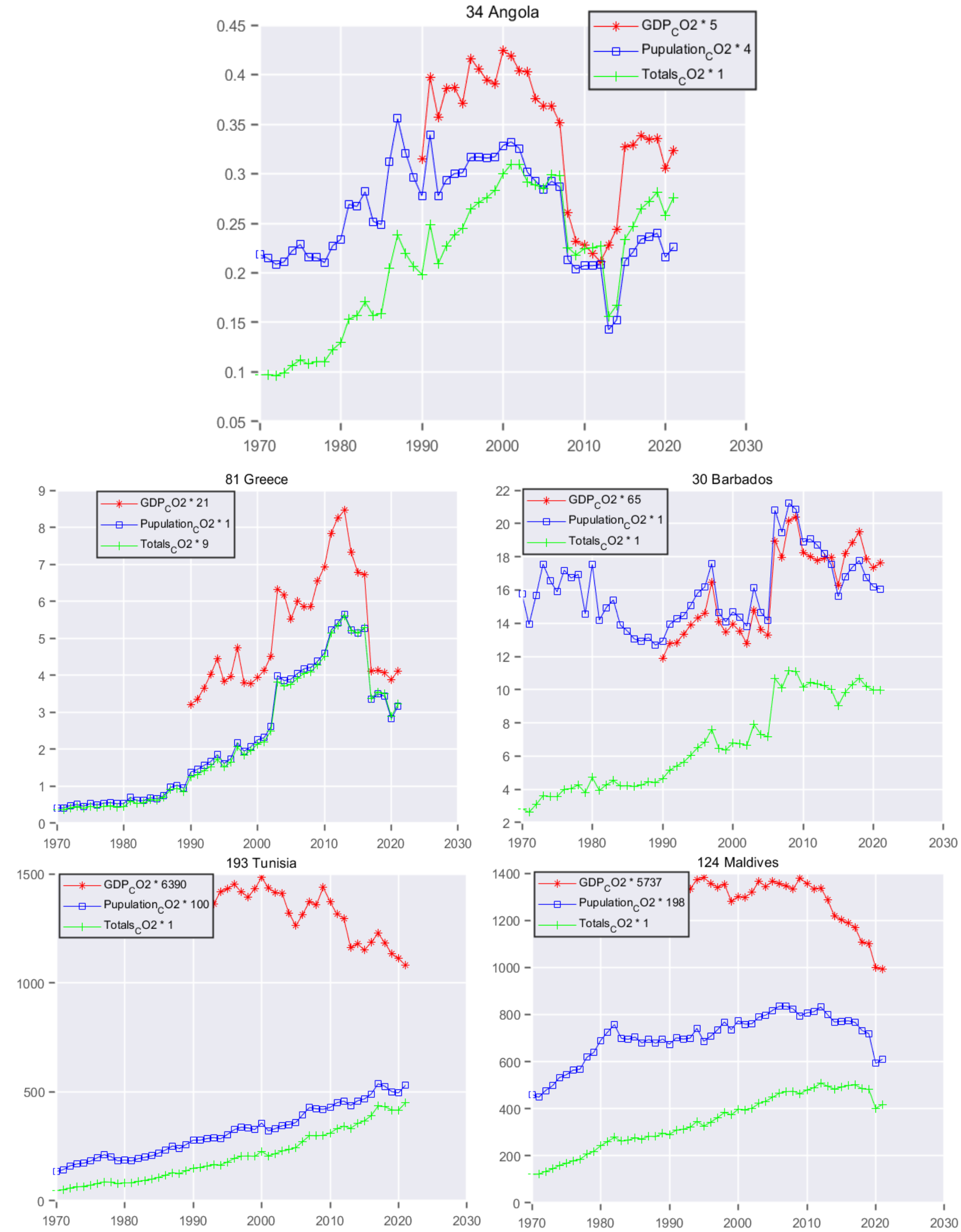


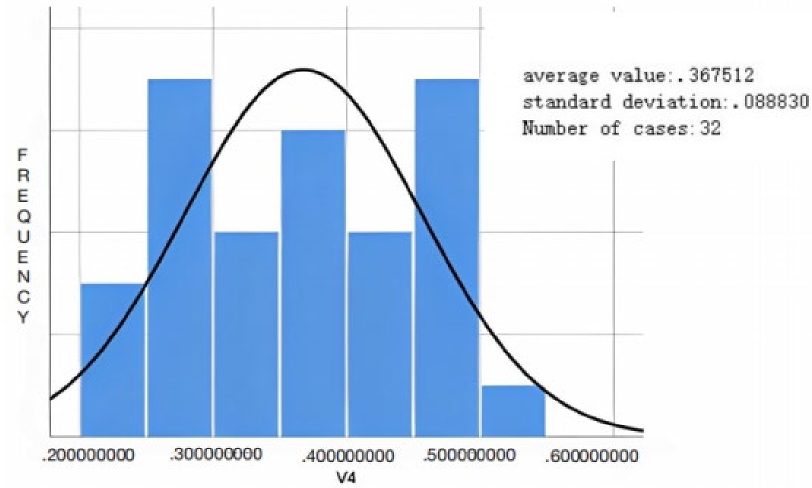
Figure 4: Change curves of three indicators

ISSN 2616-5775 Vol. 6, Issue 7: 51-58, DOI: 10.25236/AJCIS.2023.060708

The fossil carbon dioxide data per GDP of each country is negatively correlated with the total carbon dioxide data of each country. The fluctuation direction between the fossil carbon dioxide data per GDP of each country and the fossil carbon dioxide data per capita of each country is consistent. The relationship between the total carbon dioxide data of each country and the fossil carbon dioxide data per capita of each country is positive correlation.

To determine which correlation coefficient to select, Shapiro-Wilk normal distribution test (W test) is carried out for three indicator data above. It's known that Pearson correlation coefficient and typical correlation analysis require that the sample data meet the demand of normal distribution[9].

The normal distribution test is shown in Figure 5.



Normality test

	K-S inspection			Shapiro-Wilk		
	Statistics	free degree	Significance	Statistics	free degree	Significance
V1	.243	32	.000	.830	32	.000

Figure 5: Normal distribution test plot

Figure 5 indicates that the sample does not meet the requirements of normal distribution and value of significance is less than 0.05. Spearman correlation coefficient is chosen to explore the correlation.

3.2 Result analysis

Define X and Y as two sets of data, and their Spearman correlation coefficient follows formula below[10-11].

$$SP_c = 1 - \frac{6\sum_{i=1}^{n} d_i^{\,2}}{n(n^2-1)} \tag{7}$$

SP_c represents Spearman correlation coefficient, and di refers to the grade difference between X_i and Y_i.

After processing data according to the formula above, Spearman correlation coefficient of indicators is calculated. By comparison, it can be concluded that the correlation of the three indicators among the five countries is roughly similar, and the classification is reasonable. The result is shown in Figure 6.

ISSN 2616-5775 Vol. 6, Issue 7: 51-58, DOI: 10.25236/AJCIS.2023.060708

Angola

P	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1.000	0.000	0.000
Population_CO2	0.000	1.000	0.000
Totals_CO2	0.000	0.000	1.000

R	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1	-0.4213	-0.532
Population_CO2	-0.4213	1	0.9125
Totals_CO2	-0.532	0.9125	1

Greece

P	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1.000	0.031	0.003
Population_CO2	0.031	1.000	0.000
Totals_CO2	0.003	0.000	1.000

R	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1	-0.3022	-0.4195
Population_CO2	-0.3022	1	0.8847
Totals_CO2	-0.4195	0.8847	1

Barbados

P	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1.000	0.014	0.008
Population_CO2	0.014	1.000	0.000
Totals_CO2	0.008	0.000	1.000

R	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1	-0.3453	-0.4221
Population_CO2	-0.3453	1	0.915
Totals_CO2	-0.4221	0.915	1

Maldives

P	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1.000	0.001	0.002
Population_CO2	0.001	1.000	0.000
Totals_CO2	0.002	0.000	1.000

R	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1	-0.449	-0.4418
Population_CO2	-0.449	1	0.9887
Totals_CO2	-0.4418	0.9887	1

Tunisia

P	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1.000	0.000	0.000
Population_CO2	0.000	1.000	0.000
Totals_CO2	0.000	0.000	1.000

R	GDP_CO2	Population_CO2	Totals_CO2
GDP_CO2	1	-0.3772	-0.512
Population_CO2	-0.3772	1	0.9929
Totals_CO2	-0.512	0.9929	1

Figure 6: Spearman correlation coefficient of indicators

Based on the results in Figure 6, in the category of Top, the total carbon dioxide data of each country, the fossil carbon dioxide data per GDP of each country and the fossil carbon dioxide data per capita of each country have the following relationship:

- The total carbon dioxide data of each country is significantly correlated with the per capita fossil carbon dioxide data of each country, with a negative correlation coefficient of about 0.939.
- The total carbon dioxide data of each country is significantly correlated with the fossil carbon dioxide data per GDP of each country, with a negative correlation coefficient of about -0.465.
- The per capita fossil carbon dioxide data of each country and the fossil carbon dioxide data per GDP of each country have a negative correlation coefficient of about -0.379.

Then five countries are randomly selected from each of the remaining four categories and Spearman correlation analysis is performed on the three indicators of 5 different countries within the same category. Correlation between the three indicators is similar within each category, and significant differences compared to other categories are observed, thereby demonstrating the rationality of the classification.

4. Conclusions

Excessive carbon emissions have become a major global issue in recent decades. Unreasonable energy consumption structure, low energy utilization efficiency, and improper deforestation have led to a surge in carbon emissions. As a result, temperature, weather patterns and sea level of the earth has changed a lot, posing serious threats to the planet's biodiversity and human civilization.

Various countries have realized the importance of carbon emissions reduction and have taken certain measures. Therefore, a comprehensive evaluation model is needed to consider the differences in carbon emissions reduction targets, the implementation progress of current carbon emissions reduction policies, and the existence of carbon emissions reduction plans among countries.

ISSN 2616-5775 Vol. 6, Issue 7: 51-58, DOI: 10.25236/AJCIS.2023.060708

The article conducts a thorough evaluation of carbon neutral targets and related data from various nations worldwide. TOPSIS based on the EWM is utilized to assess the performance of countries' carbon emissions reduction. Based on these scores, the K-Means clustering algorithm is used to categorize the participating countries into five groups, namely Top, Good, Middle, Underachievers, and Poor. 5 countries are included in category of Top: Benin, Madagascar, Gabon, Suriname, and Georgia. Category of Good has 66 countries, represented by: Japan, France, Denmark, Tunisia, etc. 51 countries exist in category of Middle, represented by Monaco, China, Thailand, etc. 36 countries are included in category of Underachievers, represented by Jordan, South Sudan, Fiji, Comoros, etc. Category of Poor has 40 members, such as Timor-Leste, Vanuatu. This categorization provides a useful overview of the performance of different countries in terms of reducing carbon emissions.

To validate the classification, the article investigates the correlation between the total carbon dioxide data of each country, the fossil carbon dioxide data per GDP of each country, and the fossil carbon dioxide data per capita of each country. In each category, five countries are selected. Spearman correlation analysis is chosen to process the three sets of data. Each category exhibits a similar correlation pattern among its three indicators, while noticeable variations of correlation are observed when compared to the other categories. Therefore, it is determined that the classification is reasonable.

Overall, the article provides insights into the progress, status, and attitudes of countries' carbon emissions reduction, which could potentially inform policymaking and guide future planning towards a more sustainable future.

References

[1] Udara Willhelm Abeydeera L H, Wadu Mesthrige J, Samarasinghalage T I. Global research on carbon emissions: A scientometric review [J]. Sustainability, 2019, 11(14): 3972.

[2] He Y, Xing Y, Zeng X, et al. Factors influencing carbon emissions from China's electricity industry: Analysis using the combination of LMDI and K-means clustering [J]. Environmental Impact Assessment Review, 2022, 93: 106724.

[3] Mohsin M, Rasheed A K, Sun H, et al. Developing low carbon economies: an aggregated composite index based on carbon emissions[J]. Sustainable Energy Technologies and Assessments, 2019, 35: 365-374.

[4] Feng H, Wang R, Zhang H. Research on Carbon Emission Characteristics of Rural Buildings Based on LMDI-LEAP Model [J]. Energies, 2022, 15(24): 9269.

[5] Zeng Ye, Jia Heping, Yang Jing, Wang Wei, Zhong Hua, Han Jinshan, Liu Dunnan. Comprehensive Contribution Degree Evaluation Method of Virtual Power Plants Based on FAHP-EWM-TOPSIS Method [J/OL]. Modern Electric Power: 1-8[2023-03-07]. DOI: 10. 19725/j. cnki. 1007-2322. 2022. 0188.

[6] Yang Jin, Wang Zhongwei, Pang Yan. Performance evaluation of camellia industry development based on entropy TOPSIS method [J]. Journal of Central South University of Forestry and Technology, 2021, 41(12): 168-177. DOI: 10. 14067/j. cnki. 1673- 923x. 2021. 12. 019.

[7] An R, Zhu G. Clustering of economic efficiency of urban energy carbon emissions based on decoupling theory [J]. Energy Reports, 2022, 8: 9569-9575.

[8] Sammouda Rachid, El Zaart Ali. An Optimized Approach for Prostate Image Segmentation Using K-Means Clustering Algorithm with Elbow Method [J]. Computational Intelligence and Neuroscience, 2021, 2021.

[9] Mishra P, Pandey C M, Singh U, et al. Descriptive statistics and normality tests for statistical data[J]. Annals of cardiac anaesthesia, 2019, 22(1): 67.

[10] Lan Wenbao, Che Chang, Tao Chengyun. Selection of monogenic components based on Spearman rank correlation with application to SAR target recognition [J]. Chinese Journal of Radio Science, 2020, 35(03): 414- 421. DOI: 10. 13443/j. cjors. 2019063001.

[11] Wang Z, Lv J, Tan Y, et al. Temporospatial variations and Spearman correlation analysis of ozone concentrations to nitrogen dioxide, sulfur dioxide, particulate matters and carbon monoxide in ambient air, China[J]. Atmospheric Pollution Research, 2019, 10(4): 1203-1210.