INNOVATIVE MECHANICAL ENGINEERING **ISSN 2812-9229 (Online)**
University of Niš, Faculty of Mechanical Engineering
VOL. 2, NO 2, 2023, PP. 45 - 56

Original scientific paper *

APPLICATION OF PRINCIPAL COMPONENT ANALYSIS IN ASSESSING ENERGY SUSTAINABILITY OF SELECTED COUNTRIES: A CASE STUDY

Željko D. Vlaović[1], Borivoj Lj. Stepanov*[1], Mladen A. Tomić[1], Miroslav V. Kljajić[1], Đorđije D. Doder[1], Zoran M. Čepić[1]

[1]University of Novi Sad, Faculty of Technical Sciences

Abstract. *Global warming is a pressing global issue, with far-reaching consequences for the environment and human society. One critical aspect of combating global warming is reaching a high level of energy sustainability. This study investigated the energy sustainability of selected countries and compared their energy sustainability using selected energy indicators. Former Yugoslav republics (Serbia, Bosnia and Herzegovina, Montenegro, North Macedonia, Slovenia and Croatia), and the 6 top ranked countries based on the Energy Trilemma Index, were used in the analysis. Those countries are Finland, Denmark, Sweden, Canada, UK, and Switzerland. In the analysis 19 selected energy indicators were obtained from various databases like Our World in Data, EIA, and UN-stats year pocket box. Principal Component Analysis was conducted, and the result produced four clusters. The first cluster comprises Serbia, Bosnia, Montenegro, and North Macedonia. The second cluster includes Slovenia, Croatia, and the United Kingdom. The third cluster consists of Switzerland, Finland, Denmark, and Sweden. The fourth cluster is represented solely by Canada. This study's findings shed light on the energy sustainability profiles of these countries, revealing commonalities and differences within and between clusters. Such analysis increases the energy sustainability landscape understanding and can provide insights into policy recommendations and best practices.*

Key words: *Energy sustainability, PCA, Energy indicators, Country sustainability*

1. INTRODUCTION

Energy drives modern societies and economies, which ensures global prosperity. Understanding the energy sustainability profiles of different countries is crucial in order to identify areas for improvement and develop effective policies. One method to achieve this is using the Principal Component Analysis (PCA). It allows the identification of key energy indicators that contribute to a country's overall sustainability. By analyzing the data and comparing the results, patterns and trends can be identified, revealing both commonalities and differences between clusters of countries. Identification of the key energy indicators can help policymakers prioritize areas for improvement and allocate resources accordingly.

*Received: October 09, 2023 / Accepted December 25, 2023.
Corresponding author: Borivoj Stepanov
Institution: University of Novi Sad, Faculty of Technical Sciences
E-mail: bstep@uns.ac.rs

Using the PCA in analyzing energy sustainability data can greatly enhance the achievement of global energy sustainability goals and the development of effective policies. This study employs PCA to analyze various energy indicators and evaluate the overall sustainability of each selected country. Also, this research aims to provide valuable insights that can inform policy recommendations and promote best practices in achieving sustainable energy systems. Energy indicators have been crucial instruments for evaluating, comprehending, and forming energy-related policies in a time of escalating environmental concerns and the demand for sustainable development [1]. These metrics include energy use, carbon emissions, variety of energy sources, and economic impact, among many others. Energy indicators are important because they enable policymakers and analysts to make better decisions and plot a sustainable course for the future.

Advanced analytical techniques are required to comprehend the complex interactions of these energy indicators, and the PCA stands out as an important statistical technique in this context. PCA is a flexible and powerful tool that allows the reduction of data dimensionality while maintaining the essential representation of data in the initial dataset. By transforming the data into a new coordinate system determined by a collection of uncorrelated variables known as Principal Components, this technique accomplishes its goal. The primary sources of variance in the data are captured by these Principal Components, making it easier to spot underlying trends and connections between the energy indicators [2]. PCA method has wide applications. Its uses can be found in clinical studies [3], financial studies [4], sports [5] and other. Additionally, there is a wide range of applications in the field of energy, such as analyzing the performance of PV systems [6], wind turbines [7] or forecasting electricity prices [8].

In order to identify a particularly useful metric for the creation, and primarily the development, of the Sustainable Energy Action Plans, authors [9] analyzed the sustainable energy situation of a community and utilized the PCA method to evaluate the energy sustainability of rural communities. In [10], authors followed a similar approach to investigate the status and progress of sustainable household energy development in China, employing a grouped PCA method. In the context of future energy planning and the exploration of alternatives for the electric energy mix, the PCA method emerges as a highly valuable analytical tool. In [11], the main finding of the authors was the prioritization of solar resources as an energy alternative for Pakistan.

2. Problem Formulation

The primary objective of this study is to employ data mining techniques, specifically the PCA algorithm, to reduce data dimensionality. The PCA is applied to a comprehensive set of energy indicators, drawing data from diverse sources, including Our World in Data [12], the UN-Stats yearbook [13], and the U.S. Energy Information Administration (EIA) [14]. This study centers on twelve countries, selected from the former Yugoslavia and those with high living standards, identified by the World Energy Trilemma Index. (ETI) [15].

The energy sustainability understanding and monitoring in every nation are inherently connected to the utilization of energy indicators. These indicators offer insight into energy consumption, energy sources, carbon footprint, and other relevant variables. Formulating effective energy policies, managing energy resources efficiently, and achieving sustainable development all depend on a comprehensive grasp of these metrics. Tab. 1 presents the selected energy indicators. Due to the unavailability of data for the year 2022 in certain instances, data from previous years was utilized.

Table 1 Selected energy indicators

Acronym	Indicator	Unit	Ref.
F1	Carbon Intensity	$kgCO_2/kWh$	[16]
F2	Energy Intensity	kWh/GDP 2011$ (PPP)	[17]
F3	Energy use per person	kWh/cap.	[18]
F4	Coal electricity per capita	kWh/cap.	[19]
F5	Solar electricity per capita	kWh/cap.	[20]
F6	Wind electricity per capita	kWh/cap.	[21]
F7	Oil electricity per capita	kWh/cap.	[22]
F8	Hydroelectricity per capita	kWh/cap.	[23]
F9	Fossil fuel electricity per capita	kWh/cap.	[24]
F10	Renewable electricity per capita	kWh/cap.	[25]
F11	Carbon intensity of electricity	gCO_2/kWh	[26]
F12	Annual CO_2 emissions per capita	$kgCO_2/cap.$	[27]
F13	Energy security	billion kWh/ MMTOE	[28]
F14	Distribution losses	billion kWh/ MMTOE	[28]
F15	Self-Sufficiency	%	[13]
F16	Renewable share in TFEC	%	[13]
F17	Gasoline Prices per Median Income	USD/Gallon	[29]; [30]
F18	Electricity Price per Median Income	kWh/USD	[29]; [30]
F19	Gasoline Price per Median Income	kWh/USD	[29]; [30]

The variety of energy indicators presented here provides a view of a country's energy landscape, considering different aspects that span economic, environmental, and societal dimensions. The first parameter is Carbon Intensity, which specifically measures the kilograms of carbon dioxide (CO_2) emitted per kilowatt-hour. This metric is used to quantify the amount of carbon dioxide (CO_2) emissions produced per unit of energy consumed (kWh). A lower value indicates a more environmentally friendly energy mix, signifying a reduction in carbon emissions and a transition to cleaner energy sources. Parameter 2, Energy Intensity, encapsulates the relationship between energy and economic activity. This indicator assesses an economy's energy efficiency by measuring the amount of primary energy required to produce one unit of the Gross Domestic Product (GDP) per kWh in terms of 2011 USD, considering the purchasing power parities (PPP). An economy that uses fewer resources in the production of economic output, indicated by a lower value, is one that is more resource-efficient. This is frequently viewed as a good sign because it decouples economic growth from energy use, which can have a positive impact on the environment. The third parameter indicates the energy usage on a per-person basis, covering various aspects such as electricity use, heating, transportation, and cooking. A more detailed understanding is further enhanced by the parameters from 4 to 8, providing information on the per-person energy consumption from various sources. These variables reveal a country's reliance on various energy sources, including coal, oil, solar, wind, and hydroelectric power, and provide insightful information about its overall energy mix. Understanding of this energy composition is essential for making policy choices, developing energy diversification plans, and addressing climate change. The discussion of electricity generation is centered around Parameter 9, specifically focusing on electricity produced using fossil fuel such as coal, oil, and natural gas. This indicator, expressed in kWh per person, provides dependence on conventional, often less environmentally friendly energy sources for generating electrical power. Parameter 10 measures the amount of electricity produced per capita from renewable energy sources like hydropower, solar, wind, geothermal, biomass, wave, and tidal sources. Parameter 11, which represents the carbon intensity, measured in grams of CO_2-equivalents emitted per kWh of electricity. This parameter provides insights into the environmental impact of electricity generation, helping assess the carbon footprint associated with the energy produced. Parameter 12 focuses on annual CO_2 emissions per person. This indicator reveals the annual total production-based CO_2 emissions per person, measured in metric tons, offering a snapshot of a country's carbon footprint. This metric provides a comprehensive view of a country's direct contributions to global greenhouse gas emissions, as it considers territorial emissions while excluding those linked to the production of traded goods. By measuring the energy intensity of economic activity, it serves as a valuable tool for assessing and addressing environmental sustainability. Parameter 13, asses the energy security by calculating the ratio of net electricity imports to primary energy consumption. This ratio offers insight into the degree to which a country depends on imported or exported electricity relative to its overall primary energy consumption. A higher percentage signifies a greater reliance on imported electricity, exposing the nation to potential price fluctuations and disruptions in the international electricity supply. Essentially, this parameter illuminates a country's vulnerability to external influences in its energy supply chain, underscoring the crucial need for establishing a secure and self-sufficient energy infrastructure. The fourteenth parameter assesses the effectiveness of energy distribution systems by calculating the ratio of Distribution losses to primary energy consumption. This parameter specifically measures losses during the transmission of electricity and other energy sources. Higher values may indicate inefficiencies in the distribution system, leading to energy loss and the potential for negative economic and environmental effects. These energy indicators offer policymakers, researchers, and stakeholders a complete toolkit to assess and analyze a country's energy landscape from different perspectives, assisting in well-informed decision-making and sustainable energy planning. The next parameter is self-sufficiency in energy, which serves as an indicator demonstrating a country's commitment to meeting its needs through internal resources, minimizing reliance on imports or external sources. To get a better understanding of a nation's energy independence, this indicator typically assesses the proportion of domestically produced energy concerning overall consumption. Parameter 16 focuses on the renewable share in the total final energy consumption. This indicator reveals the contribution of renewable sources such as solar, wind, hydro, and biomass to the total final energy consumption. Parameters from 17 to 19 – the gasoline prices per median income, the electricity price per median income, and the gasoline price per median income - measure fuel prices relative to the median income of a country's residents. The rise in energy costs relative to income can signal potential economic stress, impacting disposable income and overall economic health. Monitoring these indicators enables informed policymaking, financial aid adjustments, or other support systems to ensure affordable and equitable energy access.

Observing energy prices concerning median income unveils potential variations in energy consumption patterns, providing insights into environmental perspectives. It reveals whether higher energy prices drive energy-saving and efficiency behaviors, ultimately reducing carbon emissions and preserving the environment.

In conclusion, the energy price indicator as a percentage of median income is pivotal for promoting energy affordability, economic health, and sustainable energy use. It assists in developing well-informed energy policies that support both a robust economy and environmental conservation. This approach aligns in-ending to reduce greenhouse gas emissions and transition toward a greener energy system by adopting cleaner and more sustainable alternatives.

Table 2 Input values of selected energy indicators and countries

Country	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
SRB	0.15	1.97	27641	3258	1.4	148.0	9.59	1204.8	3532.6	1477.6
BIH	0.18	1.77	23013	3088	21.4	116.2	30.93	1728.8	3120.4	1901.9
CRO	0.18	1.16	23468	360	36.9	507.4	2.48	1327.4	1285.2	2218.2
SLO	0.16	1.36	34391	1817	212.3	4.7	33.02	1504.8	1769.0	1853.9
MKD	0.24	1.06	13875	1041	14.3	47.5	66.87	802.4	2431.2	907.5
MNE	0.15	1.15	18940	2294	0.0	509.7	0.00	2184.7	2455.8	2695.0
SWE	0.06	1.35	59927	1	146.2	2602.4	261.63	6555.9	290.1	11128.6
DEN	0.16	0.73	32198	746	223.8	2741.6	147.90	3.4	938.4	4913.1
FIN	0.12	1.56	58966	482	54.2	1537.2	644.32	2479.8	1418.6	7231.9
UK	0.17	0.88	30098	88	185.5	966.4	137.61	84.6	2104.5	1994.7
CAN	0.14	2.44	102160	965	133.9	922.8	79.31	10207.2	2883.4	11578.9
SWE	0.12	0.60	33351	0	308.9	6.9	258.57	3902.5	258.6	4231.0

Table 3 Continued table

Coun.	F11	F12	F13	F14	F15	F16	F17	F18	F19
SRB	569	4.23	0.038	0.254	67.3	22.1	0.43	0.0068	0.0023
BIH	517	4.15	-0.735	0.179	74.7	37.1	0.31	0.0051	0.0024
CRO	246	4.36	0.338	0.194	45.5	31.6	0.22	0.0054	0.0017
SLO	237	5.92	-0.040	0.125	50.7	20.9	0.13	0.0048	0.0019
MKD	529	3.26	0.882	0.349	41.8	20.1	0.44	0.0081	0.0105
MNE	399	2.79	0.077	0.476	68.2	38.5	0.34	0.0623	0.0021
SWE	45.1	3.42	-0.468	0.164	75.6	52	0.14	0.0053	0.0047
DEN	180	5.05	0.323	0.054	76.5	37.3	0.14	0.0085	0.0035
FIN	131	6.79	0.638	0.119	57.4	45.3	0.15	0.0043	0.0037
UK	257	5.15	0.107	0.113	71.3	12.2	0.16	0.0071	0.0022
CAN	127	14.30	-0.138	0.120	175.4	22	0.10	0.0020	0.0005
SWE	45.6	4.02	0.189	0.318	53.2	25	0.11	0.0034	0.0016

To compare energy sustainability in the Western Balkan region, twelve different countries were selected for the analysis. The first six countries chosen are from the former Yugoslavia (Serbia, Bosnia and Herzegovina, Croatia, Slovenia, Montenegro and North Macedonia), while the other remaining six countries were chosen based on the World Energy Trilemma Index [31], namely: Switzerland, Canada, Sweden, Denmark, Finland and Great Britain. The previous two tables (Tab. 2. and 3.) show the input values of energy indicators for the selected countries.

3. METHODOLOGY

The initial phase of the PCA involves standardizing the input data, ensuring that the scaled dataset possesses a mean of 0 and a standard deviation of 1. This standardization is imperative to render PCA independent of variable magnitudes, thereby facilitating a more robust analysis. Subsequently, the construction of a correlation matrix using the original input data becomes pivotal. This matrix reveals how different variables are related to each other. The eigenvectors are the directions along which the data vary the most. Each eigenvector corresponds to a principal component and represents the magnitude of variance captured by each principal component. Based on eigenvalues, Principal Components are arranged in decreasing order and presented in Tab. 5. In the final step, the original data is projected into the newly created subspace that this matrix has defined. The Principal Components are captured in the resulting dataset in reduced dimensionality, providing a more condensed representation without sacrificing crucial details regarding the variability of the original data.

4. RESULTS

The PCA analysis was performed in MATLAB R2015b. Results are presented in the tables below. The correlation matrix is derived from the covariance matrix and is used to understand the relationships between different variables in the dataset.

Tab. 4 shows a 19-variable correlation matrix. The correlation matrix illuminates these variables' relationships. The matrix entries represent two variables' correlation coefficients, ranging from -1 to 1. The correlation matrix helps visualize the variable relationships. The correlation coefficient near 1 indicates a strong positive correlation, meaning that when one variable rises, the other rises too. Conversely, a coefficient nearing -1 implies a substantial negative correlation, denoting that as one variable rises, the other tends to fall. Coefficients near zero suggest the absence of a linear relationship between variables.

Looking at the correlation matrix (Tab. 4), patterns and dependencies among the 19 variables can be seen. In order to properly interpret the results shown in Tab. 4, both the magnitude and sign of the coefficients must be taken into consideration. Negative correlations indicate that variables change in opposing directions, while positive correlations indicate that variables change in the same direction. Furthermore, the degree of the correlation, approaching either 1 or -1 sheds light on how strongly the variables are related. Eigenvalues, individual and cumulative variability percentages are shown in Fig 1.

The scree plot shows the key statistical metrics. The initial two principal components cover the majority of data variance 60.7% (Fig 1. a). Post the fourth principal component, the cumulative percentage surpasses 83% (Fig. 1b). This observation implies that by including the first four Principal Components, a significant portion of the dataset's variability is effectively captured. The cumulative percentages show how efficient dimensionality reduction and data visualization are when a certain variability threshold is reached.

Table 4 Correlation matrix

1																		
-0.087	1																	
-0.535	0.597	1																
0.308	0.493	-0.304	1															
-0.354	-0.456	0.231	-0.563	1														
-0.544	-0.135	0.385	-0.471	0.247	1													
-0.518	-0.100	0.374	-0.510	0.211	0.449	1												
-0.546	0.557	0.873	-0.218	0.156	0.178	0.143	1											
0.532	0.652	-0.069	0.789	-0.646	-0.508	-0.493	-0.073	1										
-0.725	0.377	0.905	-0.439	0.249	0.633	0.469	0.858	-0.339	1									
0.694	0.219	-0.600	0.795	-0.720	-0.542	-0.562	-0.503	0.800	-0.704	1								
-0.099	0.640	0.863	-0.121	0.169	0.088	0.106	0.680	0.232	0.597	-0.352	1							
0.409	-0.397	-0.223	-0.356	-0.071	-0.048	0.289	-0.372	-0.152	-0.252	0.023	-0.086	1						
0.180	-0.186	-0.465	0.288	-0.387	-0.519	-0.295	-0.085	0.179	-0.364	0.422	-0.474	0.181	1					
-0.224	0.656	0.820	0.018	0.087	0.219	-0.091	0.776	0.306	0.675	-0.242	0.861	-0.380	-0.300	1				
-0.604	0.012	0.147	-0.078	-0.189	0.605	0.457	0.182	-0.413	0.468	-0.290	-0.229	-0.228	-0.035	-0.078	1			
0.565	0.121	-0.610	0.659	-0.775	-0.464	-0.449	-0.431	0.635	-0.608	0.929	-0.485	0.169	0.622	-0.339	-0.138	1		
0.016	-0.163	-0.321	0.311	-0.355	-0.083	-0.263	-0.138	0.178	-0.200	0.265	-0.324	0.022	0.697	-0.086	0.205	0.352	1	
0.382	-0.253	-0.313	-0.118	-0.277	0.075	0.114	-0.260	-0.042	-0.159	0.315	-0.385	0.495	0.233	-0.393	0.061	0.481	-0.062	1

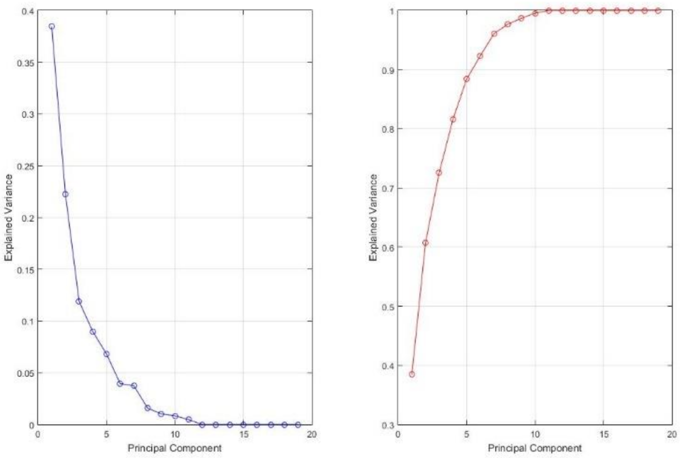

Fig. 1 Scree plot a) explained variance by single principal component b) explained variance by principal components data variance.

The dataset's PCA factor loadings are shown in Tab. 5. These loadings show how F1 to F19 and the identified principal components are related (PC1 to PC4). Positive and negative loadings with absolute values greater than 0.25 are highlighted in grey and red, respectively. Absolute values below 0.25, when not highlighted, do not exhibit significant correlations with the principal components. Overall, these weaker correlations indicate that these variables have a limited impact on the primary patterns identified by the principal components. Nevertheless, their existence suggests the possibility of nuanced connections that could be significant in some circumstances.

Table 5 Extracted principal components

	PC1	PC2	PC3	PC4
F1	-0.2728	0.054393	0.330913	0.200374
F2	0.063517	0.443381	-0.08214	0.164062
F3	0.31629	0.220167	0.00732	0.158766
F4	-0.22308	0.293278	-0.15313	-0.1442
F5	0.20883	-0.19922	0.311914	-0.30862
F6	0.230733	-0.14753	-0.22014	0.156185
F7	0.199774	-0.19752	-0.11829	0.31702
F8	0.263089	0.238064	-0.1131	0.026973
F9	-0.19622	0.38884	0.044416	0.097212
F10	0.3334	0.094389	-0.18983	0.148415
F11	-0.33212	0.175916	-0.04212	0.099619
F12	0.223488	0.303625	0.240587	0.14957
F13	-0.08645	-0.22132	0.179824	0.459729
F14	-0.21725	-0.01411	-0.26783	-0.09252
F15	0.208295	0.355142	0.042509	0.019057
F16	0.122114	-0.10934	-0.5534	0.061452
F17	-0.32552	0.10208	-0.17338	0.190534
F18	-0.13999	0.008728	-0.37844	-0.21908
F19	-0.13166	-0.16681	-0.06813	0.54217

Factor loadings are crucial to understanding relationships between original variables and Principal Components in PCA. The magnitude and direction of these loadings determine the factors' interpretation (Fig. 2). Larger absolute values denote a greater variable influence on the factor. Notably, factors F3 (0.316), F8 (0.263), F10 (0.333) are strongly positively correlated with PC1, along with substantial correlations with (F12 and F15). Conversely, they display negative correlations with F11 (loading -0.332) and F17 (loading -0.325). PC1 accommodates the greatest data variation, capturing 38.45% of the total variance.

PC2 is mostly influenced by F2 (0.443), F4 (0.293), F9 (0.388) and F15 (0.355). While no strong negative correlations are observed, F5, F6, F7, F13, and F19 exhibit approximately -0.16 loadings. PC2, the second-highest contributor, elucidates 22.25 percent of the variance. PC3 reveals three negatively correlated variables, F14, F16, and F18 - alongside positive correlations between F1 and F25, with loadings of 0.33 and 0.311, respectively. PC4 is characterized by F7, F13, and F19 as main components, featuring one negative correlation with F5. This component sheds light on specific variable relationships.

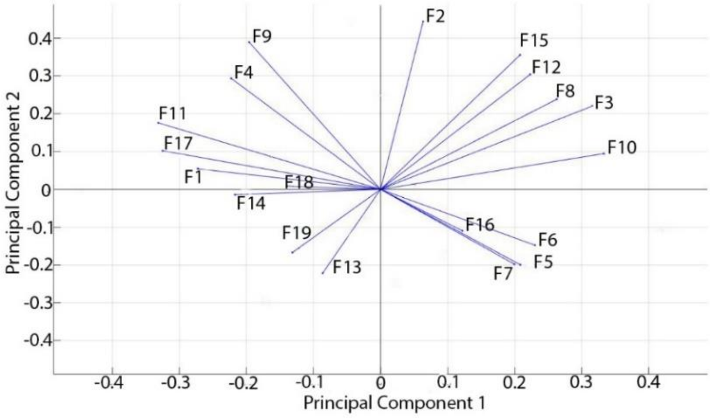

Fig. 2 Biplot

The analysis reveals that F2, representing energy intensity, exhibits significant correlations with various energy-related factors. Particularly, F2 positively correlates with variables such as annual carbon emissions per person (F12), energy use per person (F3), fossil fuel-based electricity consumption (F9), and coal electricity consumption (F4). These associations suggest that regions with higher carbon emissions per person tend to have greater energy consumption, greater reliance on fossil fuels, and higher energy intensity. Also, F2 demonstrates a positive correlation with renewable electricity consumption (F10) and hydroelectricity consumption (F8), indicating that regions embracing renewable energy sources tend to have lower carbon emissions per person. These insights emphasize the intricate link between energy consumption patterns, energy sources, and carbon emissions, underscoring the importance of sustainable energy practices in mitigating carbon footprints and promoting environmental responsibility.

The correlation between gasoline and gas prices relative to median income is crucial in energy intensity analysis. This correlation provides insights into how energy prices impact population finances (Correlation of F1 with F17, F18, and F19). The K-means algorithm was utilized to form clusters by grouping countries according to reduced dimension data as input values. The K-means algorithm detects cluster centroids in the space with reduced dimensions. The user specifies the desired number of clusters, denoted as K, which is set to 4 in this particular case. The algorithm categorizes each country into the cluster that has the closest centroid to it in the reduced-dimensional space. Figure 3 displays a Scores plot, in which each point represents a country in a space with reduced dimensions. The points are colored according to their respective clusters. In addition, the positions of the cluster centroids are indicated by the red “X” markers.

The scores plot (see Fig. 3) serves as a basis for the identification of four distinct clusters. The first cluster comprises Serbia, Bosnia and Herzegovina, Montenegro, and North Macedonia. The second cluster encompasses Slovenia, Croatia, and the United Kingdom. The third cluster includes Finland, Denmark, Sweden, and Switzerland, while Canada forms a separate cluster. Canada stands out prominently due to six energy indicators (F2, F3, F8, F10, F12, F15) manifesting maximum values, elucidating its distinctive position within the dataset."

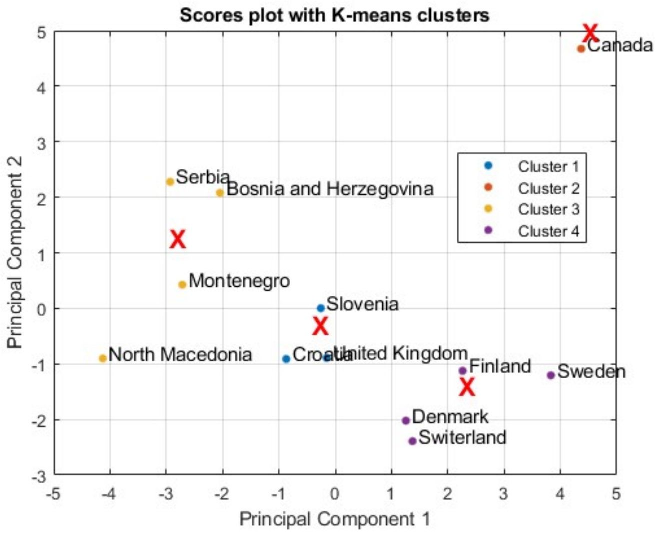

Fig. 3 Scores plot

These indicators, encompassing energy intensity, energy use per person, and electricity generation from renewable sources, collectively underscore Canada's distinctive position in the dataset. Particularly noteworthy is Canada's attainment of the highest recorded value for CO_2 emissions among the selected group of countries. Additionally, Canada has the highest level of self-sufficiency. In contrast, Switzerland ranks second lowest in six parameters (F1, F6, F11, F21, F22, and F23) and holds the lowest values for two parameters (F2 and F9). This comparison highlights the differences in energy profiles and sustainability metrics among the studied countries, providing a detailed understanding of their energy landscapes. The majority of energy-sustainable nations can be found in cluster 3, which has the lowest values for metrics like carbon intensity and the amount of electricity produced using fossil fuels. These countries rely heavily on renewable energy resources. For instance, Sweden gets 44% of its electricity from hydroelectricity [32], while Denmark gets 44% of its electricity from wind energy [33].

5. Conclusions

Energy indicators are important in today’s society, especially with environmental concerns and sustainable development. They cover energy consumption, carbon emissions, energy source diversity, and economic impact. By reducing data dimensionality while preserving the original dataset, PCA creates a new coordinate system defined by uncorrelated variables called Principal Components. These components identify the main data variances, making it easier to find energy indicator trends and connections. PCA analysis was conducted, and results were used to compare energy sustainability of 12 selected countries based on 19 selected energy indicators. Energy indicators were analyzed using data from various databases.

Significant correlations exist between annual CO_2 emissions per person (F12) and energy-related factors. Positive correlations with energy intensity (F2), energy use per person (F3), fossil fuel-based electricity consumption (F9) and coal electricity consumption (F4). Regions with higher carbon emissions per person consume more energy and use fossil fuels to generate electricity. Instead, F12 has positive correlations with renewable electricity consumption (F10) and hydroelectricity consumption (F8), indicating that regions that use renewable energy have lower carbon emissions per person. The complex relationship between energy consumption, energy sources, and carbon emissions emphasizes the need for sustainable energy practices.

Country cluster analysis shows four energy-profile-specific clusters that were made based on the K-means algorithm. First cluster countries that use fossil fuels for energy generation have high carbon emissions and energy intensity. Third-cluster countries prioritize renewable energy, reducing carbon emissions. The presence of clusters shows the global transition from fossil fuels to cleaner, more sustainable energy sources, driven by environmental, economic, and technological concerns.

REFERENCES

1. "A New World the Geopolitics of the Energy Transformation," 2019.
2. I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," The philosophical transactions, p. 374, 2016.
3. Z. Zhang and A. Castello, "Principal components analysis in clinical studies," Annals of Translational Medicine, vol. 5, p. 17, 2017.
4. N. Puri, H. Singh, A. Chakrabarti and D. Tandon, "Application of principal component analysis to facilitate financial statement analysis: an automobile industry case," Academy of Accounting and Financial Studies Journal, vol. 26, no. 1, pp. 1-19, 2022.
5. H. Weiwei, "Classification of sport actions using principal component analysis and random forest based on three-dimensional data," Displays, p. 72, 2022.
6. H. Lofti, M. Adar, A. Bennouna, D. Izbaim and F. Oumbark, "Silicon Photovoltaic Systems Performance Assessment Using the Principal Component Analysis Technique," Materials Today: Proceedings, vol. 51, pp. 1966-1974, 2022.
7. J.-c. Chu, L. Yuan, F. Xie, X.-d. Wang and L.-z. Zhang, "Operational State Analysis of Wind Turbines Based on SCADA Data," Advances in Engineering Research, vol. 185, 2019.
8. L. Burgas, J. Melendez and J. Colomer, "Principal component analysis for monitoring electrical consumption of academic buildings," Energy Procedia, vol. 62, pp. 555-564, 2014.
9. H. Doukas, A. Papadopoulou, N. Savvakis, T. Tsoutsos and J. Psarras, "Assessing energy sustainability of rural communities using Principal Component Analysis," Renewable and Sustainable Energy Reviews, vol. 16, pp. 1945-1957, 2012.
10. M. Zhang and B. Su, "Assessing Chinas rural household energy sustainable development using improved grouped principal component method," Energy, vol. 113, pp. 509-514, 2016.
11. A. Mehmood, M. Iqbal, L. Zhang, H. Xiao and J. Ren, "An application of dimensional reduction approach for evaluating energy generation alternatives for energy sustainability," Energy Reports, vol. 8, pp. 605-614, 2022.
12. https://ourworldindata.org/. (last access 12. 9. 2023.).
13. https://unstats.un.org/unsd/energystats/pubs/documents/2022pb-web.pdf. (Last access 8. 9. 2023.).
14. https://www.eia.gov/. (last access 11. 9. 2023.).
15. A.A.M. H. Al Asbahi, F. Z. Gang, W. Iqbal, Q. Abass, M. Mohsin and R. Iram, "Novel approach of Principal Component Analysis method to assess the national energy performance via Energy Trilemma Index," Energy Reports, vol. 5, pp. 704-713, 2019.
16. https://ourworldindata.org/grapher/carbon-emission-intensity-vs-gdp-per-capita.
17. https://ourworldindata.org/grapher/energy-intensity. (last access 10. 9. 2023.).
18. https://ourworldindata.org/grapher/per-capita-energy-use. (last access 11. 9. 2023.).
19. https://ourworldindata.org/grapher/coal-electricity-per-capita?tab=table. (last access 11. 9. 2023.).
20. https://ourworldindata.org/grapher/solar-electricity-per-capita. (last access 11. 9.2023.).

21. https://ourworldindata.org/grapher/wind-electricity-per-capita. (last access 11. 9. 2023.).
22. https://ourworldindata.org/grapher/oil-electricity-per-capita. (last access 11. 9. 2023.).
23. Available: https://ourworldindata.org/grapher/hydro-electricity-per-capita. (last access 11. 9. 2023.).
24. Available: https://ourworldindata.org/grapher/fossil-electricity-per-capita. (last access 12. 9. 2023.).
25. Available: https://ourworldindata.org/grapher/renewable-electricity-per-capita. (last access 12. 9. 2023.).
26. "Our World in Data - Carbon intensity of electricity," (Online). Available: https://ourworldindata.org/grapher/carbon-intensity-electricity. (Last access 13. 9. 2023.).
27. https://ourworldindata.org/grapher/co-emissions-per-capita. (Last access 9. 9. 2023.).
28. https://www.eia.gov/international/data/world/total-energy/total-energy-consumption?pd=2&p=0000002000002000020007vo70400fvu2&u=0&f=A&v=mapbubble&a=-&i=none&vo=value&t=C&g=00000000000000000000000000000000000000000000000001&l=249-ruvvvvvfvtvnvv1vrvvvvfvvvvvv.
29. https://ourworldindata.org/grapher/daily-median-income?time=latest. (last access 13. 9. 2023.).
30. https://www.visualcapitalist.com/mapped-global-energy-prices-by-country-in-2022/. (last access 14. 9. 2023.).
31. https://www.worldenergy.org/transition-toolkit/world-energy-trilemma-index. (last access 6 12 2023).
32. https://ourfuture.energy/energy-mix/sweden/. (last access 15. 9. 2023.).
33. https://www.statista.com/statistics/1235360/denmark-distribution-of-electricity-production-by-source/.(last access 11. 9. 2023.).