Sustainable Futures 10 (2025) 100907

Contents lists available at ScienceDirect

Sustainable Futures

journal homepage: www.sciencedirect.com/journal/sustainable-futures

Global development patterns: A clustering analysis of economic, social and environmental indicators

Carolina Saraiva [a], Jorge Caiado [a,b,*]

[a] ISEG Lisbon School of Management and Economics, Universidade de Lisboa, Portugal
[b] REM/CEMAPRE, Universidade de Lisboa, Portugal

ARTICLE INFO

Keywords:
Economic development
Environmental quality
Social indicators
Cluster analysis
Global development
Principal component analysis

ABSTRACT

Understanding how countries differ in their environmental and social characteristics is critical for designing effective and targeted sustainability policies. This study aims to explore heterogeneity among countries using a data-driven approach that considers multiple indicators relevant to environmental, social and economic development. We apply K-means clustering and principal component analysis (PCA) to a dataset covering 206 countries and including variables such as CO_2 emissions, access to electricity, life expectancy, education expenditure, inflation, GDP, unemployment and democracy levels. The analysis identifies eight distinct clusters of countries that broadly reflect global disparities in socio-economic and environmental conditions, with clear regional and income-based patterns. Notably, high-income democracies form distinct groups with strong social indicators, while clusters of Sub-Saharan African countries tend to exhibit low access to electricity and education spending. These findings underscore the need for differentiated policy strategies that align with each cluster's characteristics. Our results contribute to the growing literature on multidimensional sustainability assessment and offer a useful typology for international environmental cooperation and policy targeting.

1. Introduction

The interplay between economic development, environmental quality, and social indicators represents a critical area of study in contemporary socio-economic and environmental research. Understanding how these factors interact can provide valuable insights into sustainable development and policy formulation. As the world faces escalating environmental challenges and socio-economic inequalities, it becomes increasingly important to investigate how different countries navigate these complex dynamics.

Historically, the relationship between economic growth and environmental degradation has been a focal point of scholarly inquiry. The Environmental Kuznets Curve (EKC) hypothesis, which was introduced by Grossman and Krueger in the 1990s, suggests an inverted-U-shaped relationship, where environmental degradation initially increases with economic growth but eventually decreases after reaching a certain income level. This hypothesis aligns with Kuznets' theory on income inequality, which posits that economic development can lead to improved environmental quality over time [1]. However, empirical support for the EKC hypothesis has been mixed, with various studies yielding differing conclusions, depending on the environmental indicators used [2]. For instance, while some research supports the EKC hypothesis using nitrogen oxide emissions, other studies question its validity when focusing on CO2 emissions [3]. Furthermore, the literature lacks a robust typology of countries that integrates both environmental and social-economic indicators using modern, data-driven methodologies. This gap calls for a rethinking of comparative sustainability assessments through the lens of multivariate classification.

This study is motivated by the need to fill this methodological and conceptual gap. Specifically, we aim to classify countries into clusters based on a comprehensive set of indicators, including CO_2 emissions, access to electricity, life expectancy, education expenditure, inflation, GDP, unemployment and democracy levels. Our primary research question is: How can countries be grouped based on their socio-economic and environmental profiles, and what insights do these groupings provide for policy and academic debates on sustainable development?

To address this question, we apply Principal Component Analysis (PCA) to reduce the dimensionality of the dataset and capture the most significant patterns across variables, followed by K-means clustering – a widely used unsupervised machine learning technique. This combination allows us to group countries objectively without imposing any prior

* Corresponding author.
E-mail address: jcaiado@iseg.ulisboa.pt (J. Caiado).

https://doi.org/10.1016/j.sftr.2025.100907
Received 30 January 2025; Received in revised form 2 June 2025; Accepted 20 June 2025
Available online 21 June 2025

assumptions about their characteristics, while retaining the complexity of the underlying data. The goal is to identify distinct clusters that reflect meaningful differences in development patterns, thereby contributing to a more nuanced understanding of global sustainability dynamics.

The objectives of this study are fourfold: (1) to compile and standardize a set of key socio, economic and environmental indicators across countries; (2) to apply PCA and K-means clustering to identify country typologies; (3) to interpret the characteristics of each cluster in relation to global patterns of development and sustainability; and (4) to discuss the policy implications of these groupings for international cooperation and sustainable development planning. The novelty of this work lies in its integration of environmental, social, and political variables in a single clustering framework and its potential to inform differentiated, context-sensitive sustainability strategies. By offering a multidimensional classification, this study enhances both academic understanding and policy relevance in the field of global sustainability assessment.

The remainder of the paper is organised as follows: Section 2 provides a comprehensive literature review, outlining key studies and theories on the relationship between economic development, environmental quality, and social indicators, including the Environmental Kuznets Curve (EKC) hypothesis. Section 3 describes the data sources, variables, and methods used to analyse the socioeconomic and environmental performance of countries. In Section 4, we detail the empirical study, which includes the clustering analysis and principal component analysis (PCA) that group countries based on their economic, social, and environmental dimensions. This section also discusses the insights derived from the analysis of the clusters. Section 5 presents the conclusions, summarising the key findings, implications for policy and practice, and potential directions for future research.

2. Literature review

The academic and scientific focus on the relationship between the environment and the economy is longstanding, although recent years have seen a heightened emphasis on sustainability and economic growth. A notable early contribution to this discourse is the Environmental Kuznets Curve (EKC) hypothesis, proposed by Grossman and Krueger in the 1990s. The EKC posits an inverted-U relationship between economic development and environmental degradation, akin to Kuznets' theory on economic growth and income inequality. Initially, environmental degradation rises with economic growth, but after reaching a certain income level, degradation begins to decline, suggesting that economic development can eventually enhance environmental quality [1].

Despite its influence, the EKC hypothesis lacks strong empirical support, with studies producing mixed results on the causal relationship between income and environmental quality. Past research can be grouped into three categories, namely those: i) supporting, ii) refuting, or iii) presenting mixed results based on the environmental indicators used [2]. For instance, Rasli et al., [3] found that using CO2 emissions as an indicator challenges the validity of the EKC, whereas using nitrogen oxide emissions supports it.

This hypothesis, regardless of its inconsistencies, has underscored the importance of studying the intersection of environment and economy. Various studies have explored these themes, such as the mutual implications of environmental quality and economic growth [4], the relationship between CO2 emissions and economic growth in the EU [5], and the reconsideration of explanations for the EKC's inverted-U shape [6]. Almarafi et al., [7] investigated the link between the Environmental Performance Index (EPI), financial development, and economic growth.

Khatun [8] applied principal component analysis (PCA) to study environmental degradation in selected African and Asian countries, concluding that Saudi Arabia, Chile, and Malaysia were best positioned in terms of environmental quality, while most African countries showed high levels of degradation. The study also highlighted the disproportionate impact of environmental degradation on low-income populations.

Gallego-Alvarez et al. [9] employed a multivariate analysis, using an HJ-biplot, to assess environmental performance, finding that health-related environmental components were most significant in American and European countries. Chen et al. [10] used PCA to evaluate the ecological quality of Chongqing's urban area, identifying land use, pollution emission intensity, pollution control, and urban green space construction as key factors. The study emphasised the need to include economic and social indicators to fully capture environmental quality.

Jain and Mohapatra [11] developed a Composite Environmental Sustainability Index (CESI) for emerging economies (1990-2020) using PCA, with Brazil, Colombia, and Chile scoring highest in 2020. These results contrasted with those of Khatun [8], likely due to differences in temporal data and variables used.

Kumari et al., [12] studied regional disparities in socio-economic and environmental systems across Indian states using PCA, which facilitated the creation of indices for social, environmental, and economic development, and a composite index reflecting overall development. The study highlighted the strong correlation between economic, social, and environmental development.

Lu and Liu [13] applied PCA to Jiangxi province, creating separate indices for economic and environmental subsystems, which allowed for an in-depth analysis of the most influential variables in each system.

Mathrani et al., [14] used a hierarchical cluster algorithm (Ward method) to cluster 45 Asian countries based on 16 SDG indicators across four dimensions – economic, social, environmental, and institutional. Economic sustainability challenges were pervasive, but Southeast and Central Asia showed relative advantages. Socially, Central and East Asian countries performed best, while environmental sustainability varied, with Afghanistan, Sri Lanka, and Yemen performing well.

Çağlar and Gürler [15] clustered 110 countries using the k-means method, finding that low-income countries, despite better results in responsible consumption and climate change mitigation, generally clustered together. European countries tended to cluster with high SDG performance, highlighting the interdependence of sustainable development and socio-economic structures.

While Mathrani et al., [14] and Çağlar and Gürler [15] applied clustering to SDG indicators, their analyses were limited to specific regions (Asia) or omitted critical environmental metrics (e.g., CO_2 damage). Similarly, Jain and Mohapatra's [11] Composite Environmental Sustainability Index (CESI) focused narrowly on emerging economies, leaving a gap for a global assessment. Our study addresses these limitations by encompassing a broader range of indicators and countries, thus enabling cross-regional comparisons.

Jabbari et al., [16] differentiated developed and developing countries into clusters based on a socio-economic development index derived from SDG indicators. Developed countries clustered together, while developing countries formed clusters with varying environmental commitments.

Studies on the European Union include Jančovič [17], who used Eurostat data to cluster EU member states based on SDG progress, and Drastichová [18], who grouped EU countries and Norway by sustainability performance, finding that Nordic countries consistently ranked highest, although they also showed some regression in sustainability.

Beyond the Environmental Kuznets Curve literature, several studies have approached environmental sustainability from alternative theoretical and empirical perspectives. For instance, the role of institutional quality [19], international trade [20], and policy frameworks [21] have been explored as key drivers of environmental performance.

Recent typology-based studies (e.g., [22,23]) cluster countries by carbon efficiency, environmental performance indices, or sustainable development metrics. These approaches provide broader context and serve as complementary lenses to the EKC hypothesis. Additionally, debates within ecological economics challenge the growth-centric view of environmental improvement, emphasizing planetary boundaries and the concept of degrowth [24,25]. Our study contributes to this literature by integrating a multidimensional perspective in classifying countries, thus bridging methodological precision with broader policy relevance.

3. Data analytics

a) Data

This study assesses the socioeconomic and environmental performance of countries using selected variables categorised into social (5 variables), economic (4 variables), and environmental (7 variables), as shown in Table 1. The data primarily corresponds to 2020, except for the Environmental Performance Index (EPI), which incorporates earlier data due to availability constraints.

Due to incomplete data, certain potentially relevant variables, such as natural disaster occurrences and green bond issuance, were excluded from the analysis to maintain data quality. The selected variables cover key dimensions such as health, education, democracy, and infrastructure for the social dimension, represented by life expectancy (LEB), education expenditure, democracy (Dem), mortality rate, and electricity access.

Economic performance is measured through GDP per capita, inflation, unemployment, and Foreign Direct Investment (FDI), while environmental performance includes variables such as greenhouse gas emissions per capita (GHG_pc), lack of coping capacity for disasters, CO2 damage, PM2.5 exposure, and electricity installed capacity, per capita, from fossil fuels (C_Fuel_pc) and solar energy (C_Solar_pc). The EPI index further aggregates multiple indicators of climate change performance and ecosystem vitality.

For per capita variables, calculations were made using population data from the World Bank, ensuring consistent measurements across countries.

Table 1
Description of the world developing indicators used in the empirical study.

Indicator	Meaning	Source
Economic Variables		
GDP	Gross Domestic Product (per capita)	IMF[1]
Inflation	Inflation (percentage change)	IMF[1]
Unemployment	Unemployment (% of the active population)	World Bank[2]
FDI	Foreign Direct Investment (net inflows, % of GDP)	World Bank[2]
Social Variables		
Mortality rate	Number of deaths per 1,000 people	World Bank[2]
LEB	Life expectancy at birth (years)	World Bank[2]
Education Exp	Government expenditure on education (% of GDP)	World Bank[2]
Eletr Access	Population with access to electricity (%)	World Bank[2]
Dem	Level of democracy (index)	Freedom House hypothesis
Environmental Variables		
GHG_pc	Greenhouse gas emissions (tons, per capita)	World Bank[2]
Coping_Capacity	Country's lack of capacity to cope with disasters (index)	IMF[3]
CO2_damage	Cost of damages caused by CO2 emissions from fossil fuel use (% of GNI)	World Bank[2]
PM	Population exposed to more than 10 micrograms/m^3 of PM2.5 particles (% estimated values)	OECD
C_Fuel_pc	Installed electricity capacity from fossil fuels (kW, per capita)	IMF[4]
C_Solar_pc	Installed electricity capacity from solar energy (kW, per capita)	IMF[4]
EPI	Environmental Performance Index	Columbia University[4]

Note: While the database lists the "Dem" variable as covering "West Bank and Gaza", the reported value specifically reflects conditions in the West Bank only.
[1]Available at: https://www.imf.org/en/Publications/WEO/weo-database/2023/October
[2]World Development Indicators. Available at: https://databank.worldbank.org/source/world-development-indicators
[3]Available at: https://climatedata.imf.org/
[4]Available at: https://sedac.ciesin.columbia.edu/data/collection/epi/sets/browse

b) Summary statistics

Descriptive statistics were calculated for the numerical variables (Table 2) and categorical variables (Table 3) of the global development indicators.

The numerical statistics in Table 2 reveal incomplete data across variables, highlighting disparities in various development indicators. Among social variables, the stark minimum value for electricity access reflects the situation in South Sudan, where only 7% of the population has electricity. Similarly, Syria's democracy score of 1 highlights to its extreme lack of democratic governance, with 50% of countries having a democracy score below 63.

In economic terms, large disparities are observed. Burundi's GDP per capita is the lowest at $259.91, while extreme wealth in other countries drives a standard deviation higher than the mean. Inflation extremes are notable, with Venezuela experiencing hyperinflation at 2,355.15%. The median inflation rate is much lower, with half of the countries having inflation rates below 2.05%. Unemployment ranges significantly, from 0.14% in Qatar to 27.85%, with 50% of countries being below 6.45%.

Foreign Direct Investment (FDI) varies sharply, with Cyprus recording negative values due to foreign disinvestment, and the Cayman Islands showing the highest inflows. Half the countries have FDI less than 1.75% of GDP.

Environmental data also reflects large gaps. Burundi's greenhouse gas emissions per capita are the lowest, while Qatar has the highest. Lack of coping capacity values are fairly symmetrical around the mean, indicating a balanced distribution of disaster preparedness, though some countries still struggle significantly. Pollution data reveals that in half of the countries, nearly the entire population is exposed to harmful PM2.5 levels above WHO guidelines [26]

In terms of energy, fossil fuel dependency dominates over solar energy, with the former showing much higher electricity installed capacity across the board. Finally, the Environmental Performance Index (EPI) shows that countries, on average, score relatively low at 46.58, suggesting limited progress toward environmental sustainability, as 50% of countries score below 44.15, and the maximum is far below the optimal score of 100.

As indicated by Table 3, the collected sample includes countries from all continents, with the African continent being the most represented, followed by Asia, while Oceania is the least represented continent. In addition to being organised by continent, the various territories are also classified by their respective regions as shown in Fig. 1, with 22 distinct regions identified. The Caribbean region stands out as having the highest number of occurrences, while the Australia and New Zealand region has the fewest, with only three records.

After analysing the variables and their descriptive statistics, it is essential to explore potential associations among them. To this end, Pearson's correlation matrix was calculated, as shown in Fig. 2, which highlights the statistically significant correlations between world developing indicators.

The Pearson correlation coefficient measures the strength and direction of the relationship between two variables, with values ranging from −1 to 1. A coefficient closer to either extreme indicates a stronger association. Typically, a correlation of 0.7 or higher (in absolute value) signals a strong association, Ratner [27]. Additionally, a statistical significance test was conducted to account for variations in the number of observations across variables, ensuring that the calculated correlations reflect genuine associations.

Fig. 2 reveals associations among most of the selected variables, except for "FDI," which shows no significant relationship with the others. Among the remaining variables, strong positive correlations include "LEB" and "Electric Access" (0.75), "GDP" (0.71) and "EPI" (0.82), GDP and EPI (0.79) and "GHG_pc" and "C_Fuel_pc" (0.85). Notable negative correlations include "LEB" and "Coping_Capacity" (−0.89), "Electric Access" and "Coping_Capacity" (−0.76), "GDP" and "Coping_Capacity"

Table 2
Summary numerical statistics for world developing indicators.

Indicators	count	mean	std	min	median	max
Economic Variables						
GDP	190	14095	19093	260	5414	117616
Inflation	190	20.50	175.50	0.01	2.05	2355.15
Unemployment	183	8.37	5.85	0.14	6.45	27.85
FDI	185	3.31	15.82	-103.16	1.75	121.76
Social Variables						
Mortality rate	184	8.07	2.90	1.22	7.45	16.90
LEB	201	72.18	7.48	52.78	72.76	85.50
Education Exp	156	4.67	1.87	1.37	4.51	13.78
Eletr Access	206	86.26	24.45	7.26	99.95	100.00
Dem	195	58.33	30.56	1.00	63.00	100.00
Environmental Variables						
GHG_pc	189	6.20	6.57	0.34	4.62	43.33
Coping_Capacity	186	4.43	1.89	0.90	4.40	9.40
CO2_damage	187	1.87	1.62	0.07	1.33	11.50
PM	185	89.53	24.97	0.18	99.93	100.00
C_Fuel_pc	188	0.57	0.78	0.00	0.35	5.20
C_Solar_pc	185	0.07	0.13	0.00	0.01	0.81
EPI	178	46.58	15.46	22.60	44.15	82.50

Table 3
Summary categorical statistics for world developing indicators.

	count	unique	top	freq
Continent	224	5	Africa	56
Region	224	22	Caribbean	23
Economic situation	207	4	High	73

(−0.74), and "Coping_Capacity" and "EPI" (−0.86). The variables "LEB" and "Coping_Capacity" exhibit the most significant associations with other variables.

Positive correlations suggest that as one variable increases, so does the correlated variable, while negative correlations indicate an inverse relationship. For instance, a rise in per capita greenhouse gas emissions is likely accompanied by an increase in fossil fuel-based electricity capacity due to their positive association.

4. Empirical study

4.1. Methodology

In this study, we adopted cluster analysis using the non-hierarchical k-means method to form homogeneous groups of clusters of countries for each dimension of analysis (social, economic, and environmental). This method involves partitioning the data into k distinct clusters by minimising the within-cluster variance. The k-means algorithm

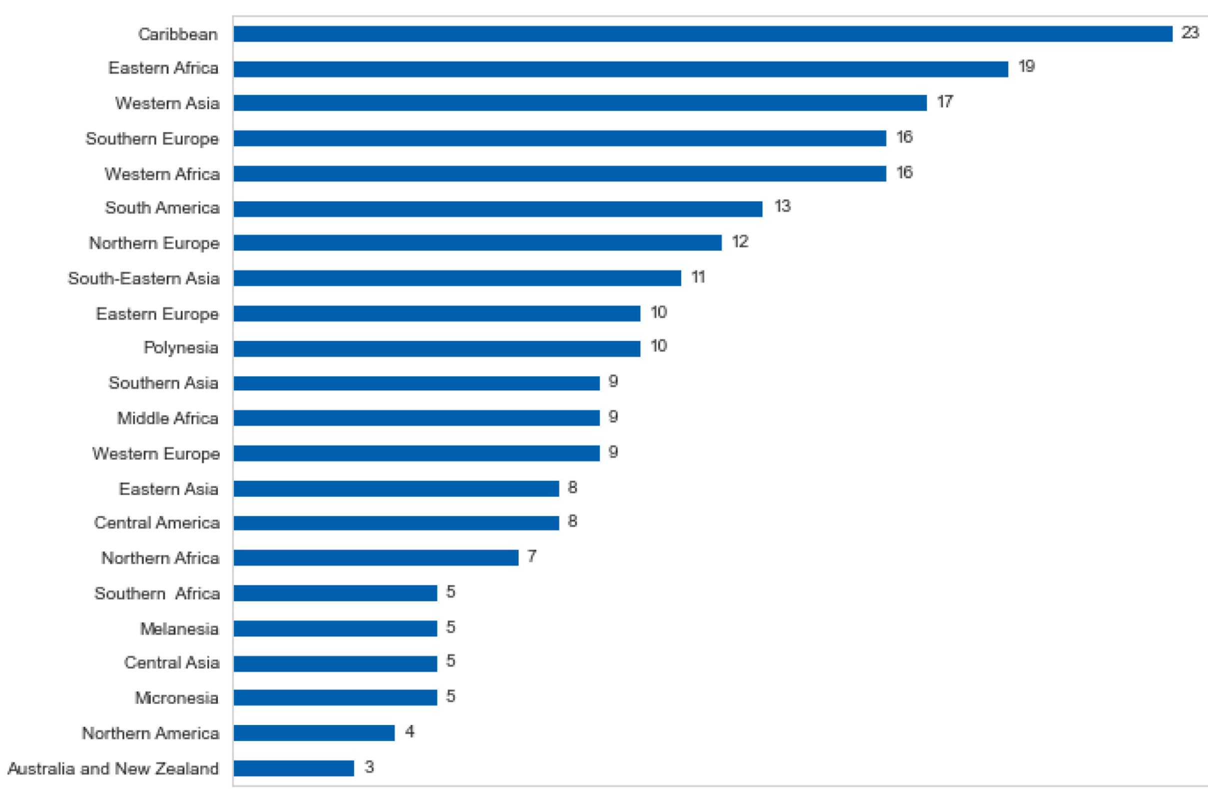

Fig. 1. Distribution of countries for region.

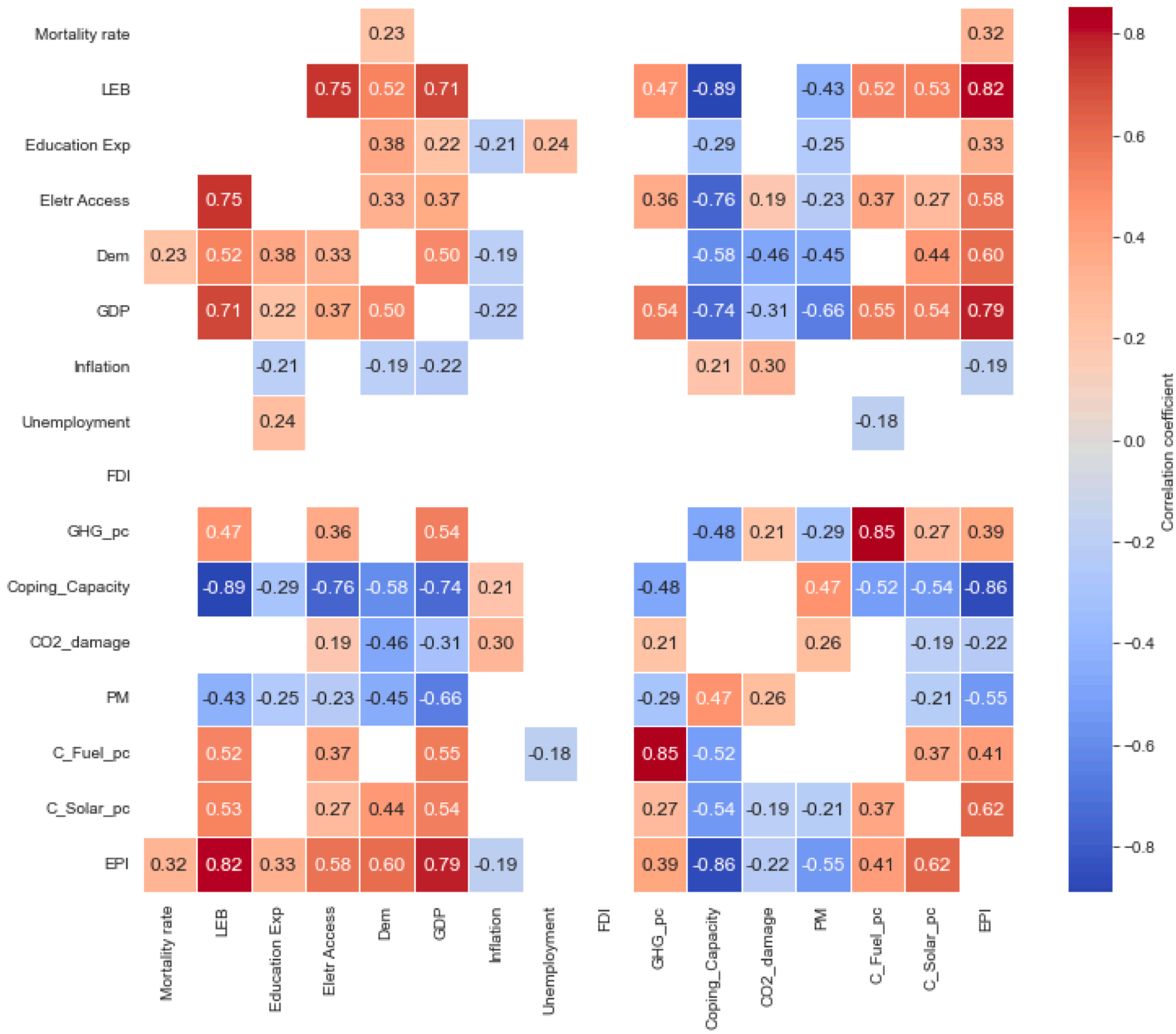

Fig. 2. Significant (p-value<0.05) Pearson correlations between world developing indicators.

iteratively assigns countries to clusters based on their proximity to the mean values (centroids) of the clusters, adjusting these centroids until convergence is achieved and the clusters are stable. For more details on these methods, refer to Hair et al., [28].

The optimal number of clusters were determined by the Silhouette Score. The Silhouette Score is a metric used to evaluate the quality of a clustering. It measures how similar an object is to its own cluster (intra-cluster similarity) compared to other clusters (inter-cluster difference), with higher values indicating better clustering. For more details, see Maharaj et al., [29].

Finally, we combined the social, economic, and environmental variables and conducted a principal component analysis (PCA) in order to reduce dimensionality and identify the underlying patterns or principal components that explain the maximum variance within the dataset. In particular, we performed PCA on standardised world developing indicators and retained components explaining over 70% of the variance [30]. These components were then used to project the countries' scores into two dimensions and visualise the optimal k-means cluster compositions.

The k-means algorithm was selected for its efficiency in partitioning large datasets into interpretable clusters, while PCA mitigates multicollinearity among variables (e.g., GDP and EPI's 0.79 correlation) and highlights dominant development drivers. This combination is optimal for answering our research questions.

Fig. 3 summarizes the modelling strategy used in this study. It outlines the main steps followed in the empirical procedure, which can be grouped into three major phases:

Step 1: Data preparation (ensure the dataset is clean, consistent and suitable for analysis), including (1) data collection and preprocessing, (2) outlier detection and removal, and (3) data standardization;
Step 2: Analytical modeling (reduce dimensionality and identify meaningful country groupings), including (4) principal component analysis (PCA) and (5) k-means clustering;
Step 3: Interpretation and application (derive insights from the clusters and link them to practical policy outcomes), including (6) interpretation of clusters and (7) policy implication discussion.

4.2. Results

a) Economic Insights

Table 4 presents a list of countries with economic variables (GDP, Inflation, Unemployment, and FDI) that have been identified as being outliers, based on a threshold of $\pm$ 3 standard deviations from the mean. By setting this threshold at ± 3 standard deviations, we are in effect isolating observations that are exceptionally far from the average, identifying them as potential outliers. These outliers may indicate countries with extreme economic conditions – either significantly better

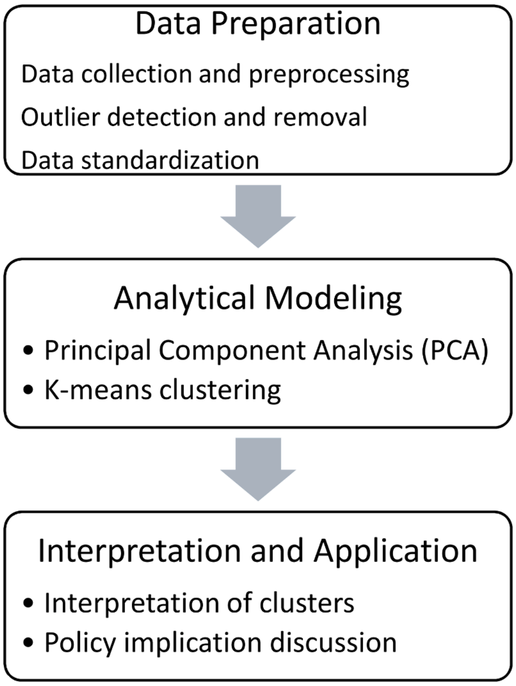

Fig. 3. Data preparation, modeling and interpretation strategy.

or worse than the global average – based on the selected variables (GDP, Inflation, Unemployment, and FDI).

High-income countries such as Ireland, Luxembourg, and Switzerland are outliers due to their extremely high GDP per capita, while low-middle-income countries such as Venezuela and Zimbabwe are outliers on account of extreme inflation, although their GDP is also very low. Countries such as Cyprus, Hungary and the Cayman Islands are outliers in FDI, with Cyprus experiencing severe disinvestment, as well as Switzerland, while Hungary and the Cayman Islands show unusually high inflows. Djibouti stands out with exceptionally high unemployment, reflecting severe socio-economic challenges.

Table 5 presents the Silhuette scores of the economic k-means clusters. The k=5 solution has the highest silhouette score (0.424), suggesting that the clustering is most effective and well-separated when using five clusters. The optimal k-means clustering results suggest different groupings of countries based on economic profiles, as shown in Table 6.

Cluster 0 includes a diverse range of countries across continents, such as the Brazil, Morocco, Greece, and South Africa. This cluster encompasses both high-income (e.g., Greece, Spain) and low-middle-income countries (e.g., Morocco, Nepal, Angola), indicating a broad mix of economic profiles. Countries in this cluster show varied economic conditions, including higher and lower GDPs, inflation rates, unemployment rates, and FDI levels.

Cluster 1 includes a large group of countries from Africa, Asia, and Latin America, such as India, Nigeria, China, and Argentina. This cluster contains a mix of emerging markets and developing economies. It shows a wide range of GDP, inflation, unemployment, and FDI levels. Countries here are generally characterised by a wider range of economic conditions, often with substantial growth potential and diverse economic challenges.

Cluster 2 consists primarily of high-income countries, including the USA, Canada, Germany, and Australia. These countries have high GDP per capita, low inflation, low unemployment, and low FDI. This cluster represents economically advanced nations with stable economic indicators. They exhibit robust economic performance and high standards of living.

Cluster 3 includes countries such as Hong Kong, Singapore, and Malta. This cluster consists of economies with high-income characteristics and small or unique economic profiles, where the countries often have high GDP per capita, low unemployment, and substantial FDI. They are economically strong, but represent unique or niche economies.

Cluster 4 includes Lebanon and Sudan. This cluster features countries with relatively lower economic indicators compared to others. Countries in this cluster face significant economic challenges, reflected in lower GDP, higher inflation, and potentially unstable economic conditions.

In short, high-income countries tend to cluster together in Cluster 2, while emerging and developing economies are spread across other clusters. Clusters with fewer countries (e.g., Cluster 4 with Lebanon and Sudan) highlight extreme cases of economic conditions.

b) Social Insights

Table 8 lists countries whose social indicators are identified as outliers based on a threshold of mean ± 3 standard deviations for the variables Mortality Rate, Life Expectancy at Birth (LEB), Education Expenditure, Access to Electricity and Democracy Index (Dem).

Serbia displays average values in most indicators, with high access to electricity and a relatively high democracy index. Cuba shows high education expenditure and access to electricity but a low democracy index. Marshall Islands and Federated States Micronesia have high democracy scores but varying education expenditure and electricity access. Burundi, Chad, Malawi, South Sudan exhibit low LEB, low

Table 5
Silhuette scores of the economic k-means clusters.

# Clusters	Score
3	0.367
4	0.388
5	**0.424**
6	0.421
7	0.331
8	0.346
9	0.355
10	0.397

Table 4
Detected and removed economic outliers using a threshold of mean ± 3 standard deviations.

Country	Continent	Region	Economic situation	GDP	Inflation	Unemploy-ment	FDI
Ireland	Europe	Northern Europe	High	**85787**	0.47	5.62	6.58
Luxembourg	Europe	Western Europe	High	**117616**	0.01	6.77	13.35
Switzerland	Europe	Western Europe	High	**86110**	0.73	4.82	-32.64
Venezuela, Rep. Bolivariana de	America	South America	Low-middle	1567	**2355.15**	7.53	NA
Zimbabwe	Africa	Eastern Africa	Low-middle	1771	**557.21**	8.65	0.70
Djibouti	Africa	Eastern Africa	Low-middle	3224	1.79	**27.85**	4.97
Cayman Islands	America	Caribbean	High	NA	NA	NA	**121.76**
Cyprus	Asia	Western Asia	High	28142	1.10	7.76	**-103.16**
Hungary	Europe	Eastern Europe	High	16094	3.33	4.25	**106.57**

Note: The detected outlier values for each numerical variable are indicated in bold. “NA” means not available.

Table 6
K-means economic cluster composition of countries.

Cluster	Country
0	'Afghanistan, Islamic Rep. of', 'Albania', 'Algeria', 'Angola', 'Armenia, Rep. of', 'Bahamas, The', 'Belize', 'Bosnia and Herzegovina', 'Botswana', 'Brazil', 'Cabo Verde', 'Chile', 'Colombia', 'Congo, Rep. of', 'Costa Rica', 'Eswatini, Kingdom of', 'Gabon', 'Georgia', 'Greece', 'Haiti', 'Honduras', 'Iraq', 'Jordan', 'Lesotho, Kingdom of', 'Mauritania, Islamic Rep. of', 'Montenegro', 'Morocco', 'Namibia', 'Nepal', 'North Macedonia, Republic of ', 'Panama', 'Rwanda', 'Samoa', 'São Tomé and Príncipe, Dem. Rep. of', 'Somalia', 'South Africa', 'Spain', 'St. Lucia', 'St. Vincent and the Grenadines', 'Tunisia'
1	'Argentina', 'Azerbaijan, Rep. of', 'Bahrain, Kingdom of', 'Bangladesh', 'Barbados', 'Belarus, Rep. of', 'Benin', 'Bhutan', 'Bolivia', 'Brunei Darussalam', 'Bulgaria', 'Burkina Faso', 'Burundi', 'Cambodia', 'Cameroon', 'Central African Rep.', 'Chad', 'China, P.R.: Mainland', 'Comoros, Union of the', 'Congo, Dem. Rep. of the', 'Croatia, Rep. of', 'Czech Rep.', 'Dominican Rep.', 'Ecuador', 'Egypt, Arab Rep. of', 'El Salvador', 'Equatorial Guinea, Rep. of', 'Estonia, Rep. of', 'Ethiopia, The Federal Dem. Rep. of', 'Fiji, Rep. of', 'Gambia, The', 'Ghana', 'Guatemala', 'Guinea', 'Guinea-Bissau', 'India', 'Indonesia', 'Iran, Islamic Rep. of', 'Jamaica', 'Kazakhstan, Rep. of', 'Kenya', 'Kyrgyz Rep.', "Lao People's Dem. Rep.", 'Latvia', 'Lithuania', 'Madagascar, Rep. of', 'Malawi', 'Malaysia', 'Maldives', 'Mali', 'Mauritius', 'Mexico', 'Moldova, Rep. of', 'Mongolia', 'Myanmar', 'Nicaragua', 'Niger', 'Nigeria', 'Oman', 'Pakistan', 'Papua New Guinea', 'Paraguay', 'Peru', 'Philippines', 'Poland, Rep. of', 'Portugal', 'Romania', 'Russian Federation', 'Saudi Arabia', 'Senegal', 'Serbia, Rep. of', 'Sierra Leone', 'Slovak Rep.', 'Solomon Islands', 'Sri Lanka', 'Suriname', 'Tajikistan, Rep. of', 'Tanzania, United Rep. of', 'Thailand', 'Togo', 'Tonga', 'Trinidad and Tobago', 'Turkmenistan', 'Uganda', 'Ukraine', 'Uruguay', 'Uzbekistan, Rep. of', 'Vanuatu', 'Vietnam', 'Zambia'
2	'Australia', 'Austria', 'Belgium', 'Canada', 'China, P.R.: Macao', 'Denmark', 'Finland', 'France', 'Germany', 'Iceland', 'Israel', 'Italy', 'Japan', 'Korea, Rep. of', 'Kuwait', 'Netherlands, The', 'New Zealand', 'Norway', 'Qatar', 'Slovenia, Rep. of', 'Sweden', 'Timor-Leste, Dem. Rep. of', 'United Arab Emirates', 'United Kingdom', 'United States'
3	'China, P.R.: Hong Kong', 'Guyana', 'Liberia', 'Malta', 'Mozambique, Rep. of', 'Singapore'
4	'Lebanon', 'Sudan'

Table 7 presents economic statistics for each of the five k-means clusters. Cluster 0 (mostly African countries) features a wide range of economic indicators, with high unemployment and fluctuating FDI and inflation. Cluster 1 (mostly African countries) includes many lower-middle-income nations with moderate inflation and unemployment, but a broad range of GDP. Cluster 2 (mainly European countries, especially from Northern Europe) consists of high-income nations with high GDP, very low inflation, low unemployment, and negative FDI. Cluster 3 (mostly Asian countries) includes high-income nations with high GDP and significant FDI, though inflation and unemployment vary. Cluster 4 (Asian countries) is characterised by extremely high inflation, high unemployment, and low GDP.

education expenditure, and very low access to electricity, with low democracy scores in some cases.

Table 9 indicates that the highest silhouette score is for 8 clusters (0.350) k-means solution. The scores are relatively close, indicating that while 8 clusters may provide the most meaningful separation, the difference is not drastic.

Table 10 shows the social cluster composition of countries of the optimal k-means solution. Table 11 presents the K-means cluster statistics for countries, based on their social indicators.

Cluster 0 includes countries such as Bolivia, Botswana, and Senegal, which exhibit diverse social indicators but share generally lower-to-middle levels in key metrics like education expenditure and mortality rate. These countries often reflect transitional development stages with mixed progress across sectors.

Cluster 1 features high-income countries such as Sweden, South Korea, and Australia, characterized by strong education systems, high life expectancy, full access to electricity, and high democracy scores. These nations are among the most socially and economically developed.

Cluster 2 includes countries such as Brazil, India, Mexico, and Algeria. This cluster is composed of upper-middle to lower-middle-income countries with a broad range of social indicators, generally reflecting moderate levels of education expenditure and access to electricity.

Cluster 3 represents countries such as Ukraine, Romania, and Georgia, primarily from Eastern Europe and post-Soviet regions. These nations show variation in democracy scores and education investment, but share historical and regional similarities.

Cluster 4 includes several Sub-Saharan African nations, such as Kenya, Ethiopia, and Rwanda. These countries generally have lower life expectancy, limited electricity access, and low education expenditure, though democracy scores vary across the group.

Cluster 5 encompasses diverse countries like Vietnam, Bangladesh, and Iran, spanning regions such as South Asia, the Middle East, and Southeast Asia. They exhibit mixed social indicators, with some showing rapid development while others maintain more modest progress in democracy and infrastructure.

Cluster 6 is composed of high-income countries like Canada, France, and Japan, known for high education expenditure, strong access to electricity, long life expectancy, and high democracy levels. The countries in this group generally enjoy strong overall social conditions.

Cluster 7 includes countries such as Namibia, Lesotho, and Sierra Leone, which face lower income levels, limited access to essential services, and higher mortality rates. Democracy indicators vary, but overall social challenges are prominent.

These clusters represent groups of countries with distinct social characteristics, driven by regional, economic, and political differences. The analysis reveals stark disparities in life expectancy, democracy, education expenditure, and access to electricity, particularly between developed regions (e.g., Europe) and less developed regions (e.g., Africa).

c) Environmental Insights

Table 12 highlights environmental outlier countries with extreme emissions, poor air quality, heavy fossil fuel use, or low resilience to disasters. High-income countries in Western Asia, like Qatar, Bahrain, and Kuwait show high emissions and poor environmental performance. In contrast, developed nations such as Norway, Germany, and New Zealand have better environmental outcomes, with lower CO2 impacts and higher sustainability scores

The environmental analysis identified 25 outliers (versus 8 social and 9 economic outliers), despite applying identical statistical criteria (mean $\pm$ 3 standard deviations). This divergence underscores two key phenomena: (1) the extreme heterogeneity in environmental performance globally, where some nations (e.g., Qatar: $GHG\_pc = 43.33$ tons) exceed others (Burundi: 0.34 tons) by orders of magnitude; and (2) the 'threshold effects' in environmental systems, where small changes in indicators like PM2.5 exposure or coping capacity can trigger disproportionate impacts [26]. Such variability is consistent with prior findings on environmental inequality [31], where resource-intensive economies cluster at one extreme and vulnerable low-income nations at the other. This contrast is less pronounced in social/economic analyses, where indicators like life expectancy or GDP exhibit more continuous distributions.

Tables 13, 14, and 15 present the results of the environmental cluster analysis. The optimal solution, with 4 clusters (Silhouette score: 0.457), distinguishes countries based on GHG emissions, environmental performance, energy sources, and lack of coping capacity.

Cluster 0 includes a diverse set of 34 countries from various continents such as Asia, Europe, and Africa. These countries tend to be upper-middle-income economies, such as China, Russia, and South Africa, which have substantial environmental challenges related to industrialisation.

Cluster 1 comprises 17 mainly European and high-income countries, including Italy, Israel, and Korea. These nations are characterised by higher per capita gas emissions and greater installed electricity capacity from both fossil fuels and solar energy.

Cluster 2 is the largest cluster, with 78 countries, mostly low- to

Table 7
K-means cluster statistics of country economic indicators.

		Cluster				
		0	**1**	**2**	**3**	**4**
	count	40	90	25	6	2
Continent	**unique**	5	5	4	4	2
	top	Africa	Africa	Europe	Asia/Africa	Asia/Africa
	freq	16	28	13	2	1
Region	**unique**	14	18	8	6	2
	top	Southern Europe	Western Africa	Northern Europe	Various regions	Western Asia/ North. Africa
	freq	6	12	6	1	1
Economic situation	**unique**	4	4	2	3	2
	top	Upper-middle	Low-middle	High	High	Upper-middle/Low
	freq	18	33	24	3	1
GDP	**mean**	6128	6410	43588	24224	2183
	std	5941	6787	14199	25757	1988
	min	611	260	1673	454	777
	max	26968	27175	68275	61274	3589
Inflation	**mean**	3.54	5.30	0.96	3.76	124.07
	std	4.95	7.15	0.77	6.55	55.42
	min	0.04	0.03	0.03	0.18	84.88
	max	22.94	42.02	2.85	16.95	163.26
Unemployment	**mean**	16.0	5.2	5.2	6.4	16.2
	std	4.0	2.4	2.3	4.6	4.2
	min	10.4	0.2	0.1	4.1	13.3
	max	24.8	11.5	9.7	15.7	19.2
FDI	**mean**	2.3	2.6	-2.6	26.1	3.9
	std	4.9	3.1	9.6	4.8	1.7
	min	-17.3	-4.9	-33.0	22.4	2.7
	max	11.2	14.0	5.8	34.1	5.1

Table 8
Detected and removed social outliers using a threshold of mean ± 3 standard deviations.

Country	**Continent**	**Region**	**Economic situation**	**Mortality rate**	**LEB**	**Education Exp**	**Eletr Access**	**Dem**
Serbia, Rep. of	Europe	Southern Europe	Upper-middle	**16.90**	74.48	3.56	100	64
Cuba	America	Caribbean	Upper-middle	10.23	77.57	**11.52**	99.93	13
Marshall Islands, Rep. of the	Oceania	Micronesia	Upper-middle	7.43	64.98	**13.78**	98.26	93
Micronesia, Federated States of	Oceania	Micronesia	Low-middle	Na	70.67	**10.54**	81.85	92
Burundi	Africa	Eastern Africa	Low	Na	61.57	5.32	**9.10**	14
Chad	Africa	Middle Africa	Low	12.49	52.78	2.74	**10.87**	17
Malawi	Africa	Eastern Africa	Low	6.71	63.72	Na	**11.50**	66
South Sudan, Rep. of	Africa	Eastern Africa	Low	10.76	55.48	Na	**7.26**	2

Note: See Table 4.

Table 9
Silhouette scores of the social k-means clusters.

# Clusters	Score
3	0.325
4	0.321
5	0.341
6	0.347
7	0.347
8	**0.350**
9	0.323
10	0.314

lower-middle-income nations from Africa, South Asia, and Central America. Countries such as Bangladesh, Brazil, and Nigeria are part of this cluster, which may indicate lower GHG emissions, but also weaker environmental infrastructure and lack of coping capacity.

Cluster 3 is composed of 12 high-income European and Latin American countries such as France, Spain, and Uruguay. This group likely represents the most environmentally advanced nations with high environmental performance indicators.

The K-means clustering analysis effectively groups countries based on their environmental characteristics, highlighting clear differences in environmental performance, emissions, and energy sources. Cluster 1 and 3, representing high-income nations, and show stronger environmental performance, but remain reliant on fossil fuels. Cluster 2, composed mostly of lower-income countries, faces significant environmental challenges related to air quality and limited environmental resilience despite low emissions. The analysis suggests that while economic status influences environmental outcomes, the global transition to cleaner energy sources, particularly solar, remains in its early stages for most countries.

d) Global Development Indicators

Fig. 4 presents the PCA loadings and explained variances of the first five principal components derived from global development indicators. PCA was applied exclusively to countries with complete data for all 16 indicators (see Appendix 1), following the removal of outliers identified in Sections 4.2. a), b) and c) (e.g., Venezuela for hyperinflation, Qatar for extreme GHG emissions). This approach ensures methodological consistency, as incomplete or outlier-distorted data can skew principal component estimation. We interpret loadings with absolute values greater than 0.4 as significant, as recommended by Hair et al., [28].

The first principal component (PC1), which explains 37.80% of the variance, is strongly associated with positive loadings for Life Expectancy at Birth (LEB), GDP per capita, Access to Electricity, Democracy Index, greenhouse gas (GHG) emissions per capita, the Environmental Performance Index (EPI), and electricity installed capacity from both fossil fuels and solar energy. These variables contribute to PC1, indicating that it reflects overall development and well-being, with a focus

Table 10
K-means social cluster composition of countries.

Cluster	Country
0	’Bolivia’, ’Botswana’, ’Nauru, Rep. of’, ’São Tomé and Príncipe, Dem. Rep. of’, ’Senegal’, ’South Africa’
1	’Australia’, ’Chile’, ’Costa Rica’, ’Cyprus’, ’Denmark’, ’Finland’, ’Iceland’, ’Israel’, ’Korea, Rep. of’, ’Malta’, ’New Zealand’, ’Norway’, ’Sweden’
2	’Algeria’, ’Bahamas, The’, ’Belize’, ’Bhutan’, ’Brazil’, ’China, P.R.: Hong Kong’, ’Colombia’, ’Dominican Rep.’, ’Ecuador’, ’El Salvador’, ’Fiji, Rep. of’, ’Ghana’, ’Honduras’, ’India’, ’Indonesia’, ’Jamaica’, ’Kyrgyz Rep.’, ’Malaysia’, ’Maldives’, ’Mexico’, ’Mongolia’, ’Morocco’, ’Nepal’, ’Panama’, ’Paraguay’, ’Peru’, ’Philippines’, ’Samoa’, ’Seychelles’, ’Suriname’, ’Tajikistan, Rep. of’, ’Timor-Leste, Dem. Rep. of’, ’Tonga’, ’Trinidad and Tobago’
3	’Armenia, Rep. of’, ’Belarus, Rep. of’, ’Croatia, Rep. of’, ’Georgia’, ’Hungary’, ’Latvia’, ’Lithuania’, ’Moldova, Rep. of’, ’Romania’, ’Russian Federation’, ’Ukraine’
4	’Cameroon’, ’Eswatini, Kingdom of’, ’Ethiopia, The Federal Dem. Rep. of’, ’Gambia, The’, ’Guinea’, ’Haiti’, ’Kenya’, ’Liberia’, ’Madagascar, Rep. of’, ’Mali’, ’Mauritania, Islamic Rep. of’, ’Niger’, ’Rwanda’, ’Tanzania, United Rep. of’, ’Togo’, ’Uganda’, ’Zambia’
5	’Azerbaijan, Rep. of’, ’Bahrain, Kingdom of’, ’Bangladesh’, ’Cambodia’, ’Comoros, Union of the’, ’Gabon’, ’Guatemala’, ’Iran, Islamic Rep. of’, ’Kazakhstan, Rep. of’, "Lao People’s Dem. Rep.", ’Lebanon’, ’Nicaragua’, ’Pakistan’, ’Qatar’, ’Singapore’, ’Thailand’, ’Turkmenistan’, ’Uzbekistan, Rep. of’, ’Vanuatu’, ’Vietnam’
6	’Albania’, ’Argentina’, ’Austria’, ’Barbados’, ’Canada’, ’Czech Rep.’, ’Dominica’, ’Estonia, Rep. of’, ’France’, ’Greece’, ’Ireland’, ’Italy’, ’Japan’, ’Luxembourg’, ’Mauritius’, ’Netherlands, The’, ’Poland, Rep. of’, ’Portugal’, ’Slovak Rep.’, ’Slovenia, Rep. of’, ’St. Vincent and the Grenadines’, ’United States’, ’Uruguay’
7	’Lesotho, Kingdom of’, ’Mozambique, Rep. of’, ’Namibia’, ’Sierra Leone’

on higher economic development, extensive infrastructure, and technological advancement.

However, the high and positive loadings of GHG emissions and fossil fuel-based electricity capacity suggest that this component also captures the environmental costs often associated with industrialisation and higher economic output. This dual nature implies that while these countries have achieved significant economic and social progress, they also face challenges related to environmental sustainability. PC1 accordingly encapsulates not just development and well-being, but also the environmental impact that often accompanies economic growth.

Given the interpretation of PC1 and its strong correlation with several key variables, PC1 can be effectively labeled as the ‘Global Development and Resilience Index’. Countries with the highest PC1 scores, such as the Republic of Korea, Israel, Austria, Italy, Singapore, France, Malta, Slovenia, and Lithuania, demonstrate the best overall levels of development and resilience. In contrast, countries with the lowest PC1 scores – such as Liberia, Niger, Guinea, Mozambique, Ethiopia, Mali, Madagascar, Haiti, Togo, Uganda, Gambia, and Rwanda – are characterised by more severe challenges in these areas.

This concept of a development index aligns with the research of Jain and Mohapatra [11], who used PCA to create a Composite Environmental Sustainability Index (CESI) for emerging economies. While their study found that Brazil, Colombia, and Chile performed similarly well, the results of our study diverge, as these countries do not rank among the highest performers. Notably, the performance gaps in our study are wider than those observed by Jain and Mohapatra. Additionally, some of the lowest-scoring countries in PC1, such as Togo, Mozambique, and Niger, align with the findings of Khatun [8], who used PCA to assess environmental degradation. This overlap suggests that these countries not only face significant environmental challenges, but also rank poorly in terms of overall global development, highlighting their persistent struggles across multiple dimensions of progress.

The second principal component (PC2, 11.35%) is strongly influenced by CO2 damage, which has a high positive loading, and moderately by GHG per capita and electricity installed capacity from fossil fuels per capita, both with positive loadings. Democracy and Education Expenses show negative loadings. This suggests that PC2 differentiates countries based on the environmental damage they cause relative to their governance structures. This component highlights the tension

Table 11
K-means cluster statistics of country social indicators.

		Cluster							
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**
	count	6	13	34	11	17	20	23	4
Continent	**unique**	3	4	4	2	2	4	4	1
	top	Africa	Europe	America	Europe	Africa	Asia	Europe	Africa
	freq	4	6	15	9	16	15	14	4
Region	**unique**	5	7	12	4	5	8	9	4
	top	Southern Africa	Northern Europe	South America	Eastern Europe	Eastern Africa	South-Eastern Asia	Southern Europe	Southern Africa
	freq	2	5	6	6	7	5	5	2
Economic situation	**unique**	3	2	3	3	2	3	2	3
	top	Low-midlle	High	Low-midlle	Upper-midlle	Low	Low-midlle	High	Low
	freq	3	12	16	6	10	10	18	2
Mortality rate	**mean**	7.36	7.25	6.67	14.83	7.73	5.96	10.19	9.66
	std	1.62	1.54	1.28	1.05	1.28	1.75	1.74	2.2
	min	5.58	5.3	3	12.79	5.85	1.22	6.5	8
	max	9.43	10	9.33	16.55	9.85	8.6	12.6	12.86
LEB	**mean**	65.77	82.01	72.63	73.44	62.63	72.45	78.98	59.61
	std	1.82	1.34	3.75	2.28	2.46	5.34	3.26	3.51
	min	63.44	79.28	64.11	70.17	58.63	64.17	72.13	54.69
	max	68.01	83.43	85.5	77.72	66.77	84.47	84.56	62.83
Education Exp	**mean**	6.64	6.55	4.78	4.42	3.29	3.09	4.72	7.93
	std	1.34	1.2	0.99	1.03	1.04	0.97	0.73	1.31
	min	5.02	4.8	2.87	2.71	1.37	1.67	3.27	6.64
	max	8.44	8.61	7.04	6.6	5.36	4.95	5.68	9.28
Eletr Access	**mean**	83.36	99.99	98.04	99.99	48.27	95.14	99.98	39.18
	std	14.51	0.03	3.22	0.03	15.35	8.34	0.11	12.65
	min	64.3	99.9	85.44	99.9	18.72	67.3	99.46	26.31
	max	100	100	100	100	79.99	100	100	52.3
Dem	**mean**	74.83	93.38	63.56	62.09	38	29	90.91	62
	std	6.43	7.17	18.85	26.37	13.49	18.62	7.3	14.09
	min	66	76	8	11	16	2	66	43
	max	84	100	91	90	60	82	98	77

Table 12
Detected and removed environmental outliers using a threshold of mean ± 3 stand. deviations.

Country	Continent	Region	Economic situation	GHG	Coping	CO2	PM	Fuel	Solar	EPI
Bahrain	Asia	Western Asia	High	**36.7**	3.0	4.10	99.9	5.20	0.007	51.0
Brunei Darussalam	Asia	South-Eastern Asia	High	**27.0**	3.5	2.42	4.6	2.78	0.003	54.8
Kuwait	Asia	Western Asia	High	**31.2**	3.7	Na	99.7	4.62	0.019	53.6
Qatar	Asia	Western Asia	High	**43.3**	2.9	2.66	100.0	3.83	0.002	37.1
Turkmenistan	Asia	Central Asia	Upper-middle	**31.1**	5.9	Na	100.0	1.12	Na	43.9
United Arab Emirates	Asia	Western Asia	High	**26.9**	1.8	2.12	99.9	3.39	0.251	55.6
Iran, Islamic Rep. of	Asia	Southern Asia	Low-middle	9.7	4.6	**11.50**	99.9	0.81	0.005	48.0
Mongolia	Asia	Eastern Asia	Low-middle	16.4	4.6	**7.53**	91.8	0.37	0.029	32.2
Syrian Arab Rep.	Asia	Western Asia	Low	2.3	5.5	**9.85**	100.0	0.42	0.001	Na
Uzbekistan, Rep. of	Asia	Central Asia	Low-middle	5.5	3.8	**7.81**	100.0	0.43	0.000	44.3
Canada	America	Northern America	High	17.8	2.4	1.41	**0.5**	0.93	0.095	71.0
Iceland	Europe	Northern Europe	High	7.2	1.9	0.29	**3.7**	0.32	0.019	72.3
Ireland	Europe	Northern Europe	High	11.9	1.8	0.46	**3.4**	1.24	0.018	72.8
Mauritius	Africa	Eastern Africa	Upper-middle	5.1	2.8	1.27	**10.5**	0.51	0.085	45.1
New Zealand	Oceania	Australia and New Zealand	High	15.8	1.8	0.67	**0.2**	0.43	0.028	71.3
Norway	Europe	Northern Europe	High	8.6	1.6	0.39	**0.8**	0.18	0.030	77.7
United States	America	Northern America	High	16.6	2.2	0.89	**8.9**	2.21	0.231	69.3
Bahamas	America	Caribbean	High	6.9	3.1	0.98	98.7	**2.16**	0.007	43.5
Oman	Asia	Western Asia	High	20.9	3.7	4.46	99.9	**2.31**	0.028	38.5
Saudi Arabia	Asia	Western Asia	High	19.8	3.4	2.93	99.8	**2.21**	0.003	44.0
Australia	Oceania	Australia and New Zealand	High	22.3	2.1	1.23	18.2	2.03	**0.812**	74.9
Belgium	Europe	Western Europe	High	8.7	1.8	0.68	86.6	0.60	**0.483**	73.3
Germany	Europe	Western Europe	High	8.3	1.6	0.65	55.9	1.03	**0.645**	77.2
Japan	Asia	Eastern Asia	High	8.7	1.5	0.82	98.4	1.53	**0.569**	75.1
Netherlands	Europe	Western Europe	High	8.9	1.3	0.66	85.7	1.31	**0.637**	75.3

Table 13
Silhuette scores of the environmental k-means clusters.

# Clusters	Score
3	0.431
4	**0.457**
5	0.416
6	0.413
7	0.405
8	0.407
9	0.368
10	0.353

between environmental impact (CO2 damage) and governance factors (democracy), which is alignment with the findings of Gültekin et al., [2], who concluded that countries with lower levels of democracy tend to experience higher levels of pollution.

The third principal component (PC3, 9.46%) has high negative loadings for Unemployment and Education Expenditure, while Foreign Direct Investment (FDI) has a strong positive loading. This component appears to capture the relationship between economic instability (unemployment, low education expenditure) and foreign investment, as well as health outcomes (mortality).

The fourth principal component (PC4, 7,37%) captures a relatively small portion of the total variance compared to PC1, but still provides meaningful insights. The contrast it highlights – between countries with higher mortality and lower unemployment versus those with lower educational expenditure and poorer air quality – can reveal important social and environmental dynamics in specific groups of countries.

Finally, the fifth principal component (PC5, which accounts for 6.81% of the variance), though explaining slightly less variance than PC4, highlights a unique economic dimension, which is particularly centered on inflation and its impacts. This component contrasts economic instability, which is chracaterised by high inflation, with more favourable life outcomes, such as moderate to high life expectancy and widespread access to electricity. PC5 thus reflects the disparities in economic and social resilience, illustrating how inflationary pressures caneither coexist with or counteract other aspects of well-being.

The k-means cluster analysis of global development indicators (Tables 16, 17 and 18) provides insights into how countries are grouped

Table 14
K-means envirnomental cluster composition of countries.

Cluster	Country
0	'Algeria', 'Argentina', 'Armenia, Rep. of', 'Azerbaijan, Rep. of', 'Belarus, Rep. of', 'Bosnia and Herzegovina', 'China, P.R.: Mainland', 'Egypt, Arab Rep. of', 'Georgia', 'Grenada', 'Guyana', 'India', 'Iraq', 'Jamaica', 'Jordan', 'Kazakhstan, Rep. of', "Lao People's Dem. Rep.", 'Lebanon', 'Malaysia', 'Maldives', 'Marshall Islands, Rep. of the', 'Mexico', 'Moldova, Rep. of', 'Montenegro', 'North Macedonia, Republic of ', 'Russian Federation', 'South Africa', 'Suriname', 'Tajikistan, Rep. of', 'Thailand', 'Trinidad and Tobago', 'Tunisia', 'Ukraine', 'Vietnam'
1	'Antigua and Barbuda', 'Barbados', 'Bulgaria', 'Chile', 'Cyprus', 'Czech Rep.', 'Greece', 'Hungary', 'Israel', 'Italy', 'Korea, Rep. of', 'Malta', 'Poland, Rep. of', 'Romania', 'Singapore', 'Slovak Rep.', 'Slovenia, Rep. of'
2	'Afghanistan, Islamic Rep. of', 'Albania', 'Angola', 'Bangladesh', 'Belize', 'Benin', 'Bhutan', 'Bolivia', 'Botswana', 'Brazil', 'Burkina Faso', 'Burundi', 'Cabo Verde', 'Cambodia', 'Cameroon', 'Central African Rep.', 'Chad', 'Colombia', 'Congo, Dem. Rep. of the', 'Congo, Rep. of', 'Costa Rica', 'Djibouti', 'Dominica', 'Dominican Rep.', 'Ecuador', 'El Salvador', 'Equatorial Guinea, Rep. of', 'Eswatini, Kingdom of', 'Ethiopia, The Federal Dem. Rep. of', 'Fiji, Rep. of', 'Gabon', 'Gambia, The', 'Ghana', 'Guatemala', 'Guinea', 'Guinea-Bissau', 'Haiti', 'Honduras', 'Indonesia', 'Kenya', 'Kiribati', 'Lesotho, Kingdom of', 'Liberia', 'Madagascar, Rep. of', 'Malawi', 'Mali', 'Mauritania, Islamic Rep. of', 'Micronesia, Federated States of', 'Morocco', 'Mozambique, Rep. of', 'Myanmar', 'Namibia', 'Nepal', 'Nicaragua', 'Niger', 'Nigeria', 'Pakistan', 'Panama', 'Papua New Guinea', 'Peru', 'Philippines', 'Rwanda', 'Samoa', 'São Tomé and Príncipe, Dem. Rep. of', 'Senegal', 'Sierra Leone', 'Solomon Islands', 'Sri Lanka', 'St. Lucia', 'St. Vincent and the Grenadines', 'Tanzania, United Rep. of', 'Timor-Leste, Dem. Rep. of', 'Togo', 'Tonga', 'Uganda', 'Vanuatu', 'Zambia', 'Zimbabwe'
3	'Austria', 'Croatia, Rep. of', 'Denmark', 'France', 'Latvia', 'Lithuania', 'Paraguay', 'Portugal', 'Spain', 'Switzerland', 'United Kingdom', 'Uruguay'

Note: Notably, Finland and Sweden were omitted from this analysis due to unavailable PM2.5 data, though their GHG emissions (8.6 and 4.4 tons/capita, respectively) and EPI (77.7 and 78.7, respectively) suggest they would likely align with Cluster 3's high-performance group. In the same way, two other European countries (Estonia and Luxembourg) could not be classified for lack of complete PM2.5 readings. Based on the indicators that are available,Estonia (GHG = 7.7, EPI = 65.3) and Luxembourg (GHG = 14.0, EPI = 82.3) resemble the high-income, high-GHG economies gathered in Cluster 1.

Table 15
K-means cluster statistics of country envirnomental indicators.

		Cluster			
		0	**1**	**2**	**3**
	count	34	17	78	12
Continent	**unique**	5	3	5	2
	top	Asia	Europe	Africa	Europe
	freq	15	10	41	10
Region	**unique**	13	7	16	4
	top	Western Asia	Eastern Europe	Western Africa	Northern Europe
	freq	6	6	15	4
Economic situation	**unique**	3	2	3	2
	top	Upper-middle	High	Low-middle	High
	freq	25	15	40	11
GHG	**mean**	6.77	8.14	2.32	6.61
	std	4.37	2.93	1.40	1.56
	min	1.85	4.04	0.34	5.09
	max	19.39	13.03	6.48	10.50
Coping	**mean**	4.26	2.41	5.67	2.17
	std	0.85	0.68	1.37	0.91
	min	2.90	1.00	2.50	0.90
	max	6.50	3.70	8.80	4.30
CO2	**mean**	3.27	1.19	1.41	0.64
	std	1.17	0.50	0.60	0.28
	min	1.19	0.41	0.27	0.20
	max	5.39	2.26	3.36	1.05
PM	**mean**	98.18	98.47	98.66	43.82
	std	5.67	3.69	4.39	20.50
	min	71.52	84.58	69.72	15.92
	max	100	100	100	79.16
Fuel	**mean**	0.63	0.99	0.11	0.49
	std	0.32	0.48	0.11	0.30
	min	0.08	0.41	0.00	0.00
	max	1.59	2.00	0.49	0.95
Solar	**mean**	0.04	0.20	0.01	0.14
	std	0.05	0.10	0.02	0.11
	min	0.00	0.06	0.00	0.00
	max	0.18	0.36	0.08	0.34
EPI	**mean**	44.2	63.1	35.6	69.1
	std	6.8	7.9	7.3	12.8
	min	27.6	45.6	22.6	46.4
	max	55.4	72.0	52.9	82.5

based on various development metrics. The optimal number of clusters appears to be 8, as indicated by the highest silhouette score.

Cluster 0 includes high-income and highly developed countries such as Israel, Korea, Malta, and Singapore. These countries likely have high development indicators.

Cluster 1 consists of a mix of middle-income countries with diverse development levels, including Argentina, India, and Mexico. This cluster represents a range of developing countries with varying degrees of economic and social development.

Cluster 2 contains low-income and developing countries, such as Ethiopia, Ghana, and Uganda. These countries are likely characterized by lower development indicators and higher economic challenges.

Cluster 3 includes mostly European countries, such as Austria, Greece, and Italy. These countries are generally high-income with advanced development indicators.

Cluster 4 features countries in transition or with emerging economies, such as Belarus, Kazakhstan, and Ukraine, all of which have intermediate development levels.

Cluster 5 includes countries with significant economic and development challenges, such as Algeria, Pakistan, and Vietnam.

Cluster 6 consists of countries with relatively higher development levels, including France, Latvia, and Uruguay.

Cluster 7 comprises countries with diverse but generally lower-middle to upper-middle economic development, such as Bolivia and Namibia.

The clusters represent a range of economic and development levels, from highly developed countries (Cluster 0) to low-income, developing countries (Cluster 2). The composition of each cluster highlights different stages of development and economic conditions, providing a nuanced view of global development indicators.

Fig. 5 presents a clustering of countries based on the two principal components derived from global development indicators. The distribution reveals clear patterns of similarity and disparity among countries across the two dimensions (PC1 and PC2).

European countries, primarily from Clusters 0, 3, and 6 – such as Malta, Austria, Greece, Italy, Slovakia, Slovenia, Croatia, France, Latvia, Lithuania, and Romania – are grouped in the positive region of PC1 and the negative region of PC2. This indicates strong economic development, favourable environmental outcomes, and higher life expectancy, all of which reflects strong performance on global development. These results align with Çağlar and Gürler [15], where high-income countries also cluster together, with top scores on the Sustainable Development Goal Index.

In contrast, countries with lower economic development, less favourable environmental outcomes, and lower life expectancy — including Algeria, Azerbaijan, Cambodia, India, Indonesia, Laos, Lebanon, Nicaragua, Pakistan, Suriname, Tajikistan, and Vietnam — are clustered together in the negative region of PC1 and the positive region of PC2. This indicates shared challenges across these nations in terms of developmental metrics.

Additionally, a distinct large cluster appears in the negative regions of both PC1 and PC2, consisting mainly of countries from Cluster 2 and parts of Cluster 1. These countries face significant economic and developmental challenges, performing the worst in PC1. This aligns with Çağlar and Gürler [15], where clusters of low and middle-low-income countries also show poorer results in the Development Index.

Finally, post-Soviet states, including Ukraine, Moldova, Belarus, Kazakhstan, and the Russian Federation, primarily from Cluster 4, appear closely grouped in another section of the map. Their proximity in this region reflects shared historical, economic, and developmental characteristics within the post-Soviet geopolitical sphere.

4.3. Policy implications of cluster findings

The clustering analysis reveals distinct development trajectories across nations, each necessitating targeted policy interventions to address their unique socio-economic and environmental challenges.

a) High-Income Clusters

For high-income clusters (Clusters 0, 3 and 6), exemplified by countries such as Israel, Singapore, and France, the primary challenge lies in decoupling economic growth from environmental degradation. Despite achieving high Environmental Performance Index (EPI) scores (65.3, 65.9 and 61.1, respectively) and robust social indicators (e.g., universal electricity access, life expectancy > 83, 79 and 76 years, respectively), these nations exhibit persistently elevated greenhouse gas (GHG) emissions (6.7–9.3 tons/capita), driven by energy-intensive industries and transportation systems. To reconcile this paradox, policymakers should prioritize carbon pricing mechanisms, such as phased fossil-fuel taxation modeled after Sweden's successful $137/ton CO_2 levy, with revenue recycling to mitigate regressive impacts on low-income households. Concurrently, scaling up distributed solar capacity, as demonstrated by Australia's outlier solar adoption (0.81 kW/capita) through feed-in tariffs, could accelerate renewable energy transitions. Furthermore, industrial symbiosis policies, akin to the EU's Ecodesign Directive, should be mandated to reduce PM2.5 exposure, which remains significantly lower in these clusters (e.g., 56.0% in Cluster 6) compared to global averages (92.7%).

b) Low-Income Clusters

In contrast, low-income clusters (Clusters 2 and 7), including Ethiopia, Niger, and Mozambique, face compounded vulnerabilities: extreme energy poverty (50% average electricity access in Cluster 2; South Sudan at 7.26%), high disaster susceptibility (Coping_Capacity

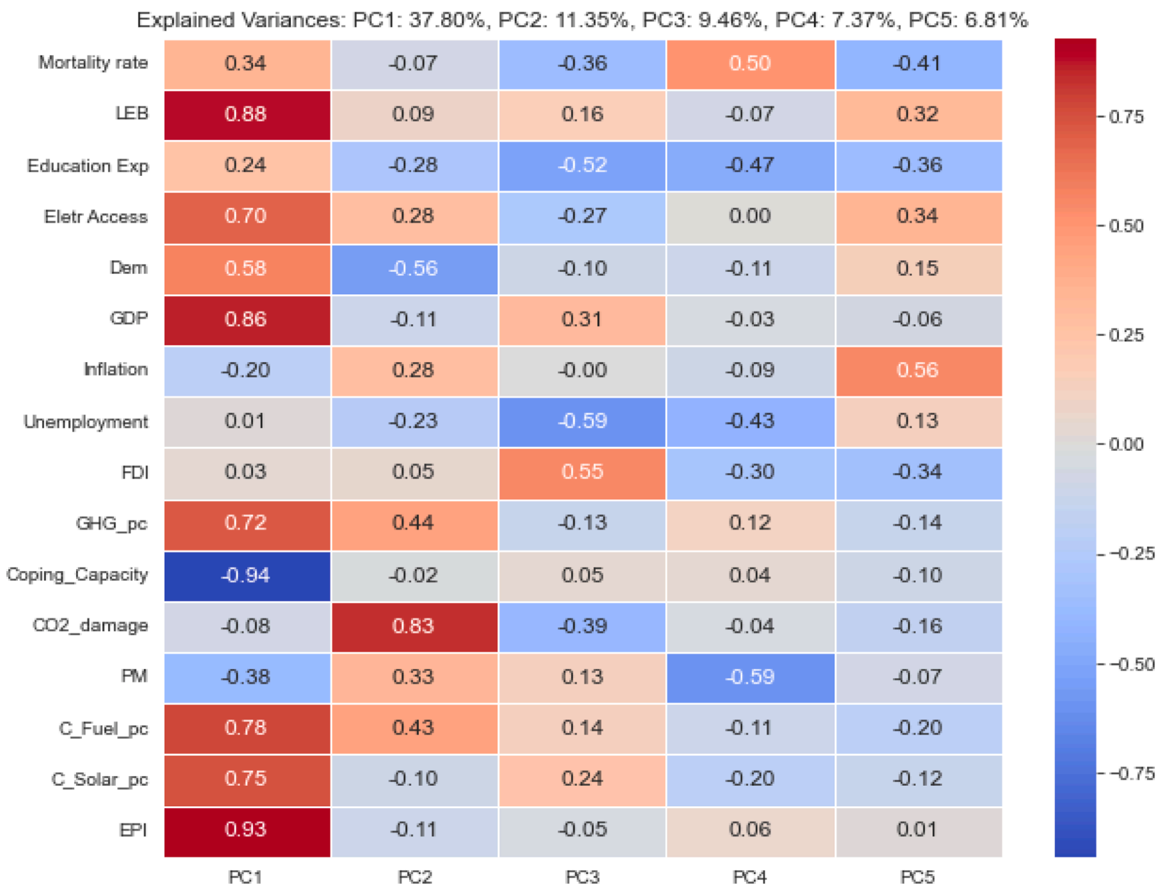

Fig. 4. PCA loadings heatmap of global development indicators.

Table 16
Silhouette scores for k-means clusters of global development indicators.

# Clusters	Score
3	0.280
4	0.289
5	0.296
6	0.287
7	0.289
8	**0.308**
9	0.239
10	0.256

Table 17
Composition of countries in k-means clusters based on global development indicators.

Cluster	Country
0	'Israel', 'Korea, Rep. of', 'Malta', 'Singapore'
1	'Albania', 'Argentina', 'Armenia, Rep. of', 'Bangladesh', 'Belize', 'Bhutan', 'Brazil', 'Colombia', 'Costa Rica', 'Dominican Rep.', 'Ecuador', 'El Salvador', 'Fiji, Rep. of', 'Gabon', 'Georgia', 'Guatemala', 'Honduras', 'India', 'Indonesia', 'Jamaica', 'Maldives', 'Mexico', 'Morocco', 'Nepal', 'Nicaragua', 'Panama', 'Peru', 'Philippines', 'Samoa', 'São Tomé and Príncipe, Dem. Rep. of', 'Timor-Leste, Dem. Rep. of', 'Tonga'
2	'Cambodia', 'Cameroon', 'Ethiopia, The Federal Dem. Rep. of', 'Gambia, The', 'Ghana', 'Guinea', 'Haiti', 'Kenya', 'Liberia', 'Madagascar, Rep. of', 'Mali', 'Mauritania, Islamic Rep. of', 'Mozambique, Rep. of', 'Niger', 'Rwanda', 'Senegal', 'Sierra Leone', 'Tanzania, United Rep. of', 'Togo', 'Uganda', 'Vanuatu', 'Zambia'
3	'Austria', 'Barbados', 'Chile', 'Czech Rep.', 'Greece', 'Italy', 'Poland, Rep. of', 'Slovak Rep.', 'Slovenia, Rep. of'
4	'Belarus, Rep. of', 'Kazakhstan, Rep. of', 'Malaysia', 'Moldova, Rep. of', 'Russian Federation', 'Trinidad and Tobago', 'Ukraine'
5	'Algeria', 'Azerbaijan, Rep. of', "Lao People's Dem. Rep.", 'Lebanon', 'Pakistan', 'Suriname', 'Tajikistan, Rep. of', 'Thailand', 'Vietnam'
6	'Croatia, Rep. of', 'France', 'Latvia', 'Lithuania', 'Paraguay', 'Romania', 'Uruguay'
7	'Bolivia', 'Botswana', 'Eswatini, Kingdom of', 'Lesotho, Kingdom of', 'Namibia', 'South Africa', 'St. Vincent and the Grenadines'

mean = 6.4 in Cluster 2), and near-universal exposure to hazardous PM2.5 levels (above 99%). Addressing these challenges requires integrated solutions that bridge basic needs and resilience. Electrification efforts must leverage off-grid solar microgrids, building on models like Kenya's pro-poor rooftop solar programs, which achieved 50% adoption through public-private partnerships. Simultaneously, community-based adaptation strategies, such as Bangladesh's cyclone early-warning systems, should be mainstreamed into national development plans to enhance disaster preparedness. Health-environment co-benefits can be realized through subsidies for clean cookstoves, which reduce indoor air pollution linked to elevated mortality rates (7.45 and 9.66 deaths/1000, respectively) in these regions.

c) Transitional Economies

Transitional economies (Cluster 4), such as Ukraine, Kazakhstan, and Russia, exhibit mixed trajectories characterized by moderate GDP/capita ($8,393) but volatile governance deficits (mean democracy score = 44). To stabilize growth and align with sustainability goals, these nations should condition FDI on green manufacturing, emulating Vietnam's solar panel export hubs through tax incentives. Anti-corruption measures, including transparency benchmarks tied to EU accession funds, as implemented in Moldova's 2021 judicial reforms, could strengthen institutional capacity. Additionally, a just transition for fossil-fuel-dependent workers is critical; Germany's Coal Phase-Out Act, which allocated €40 billion for retraining and regional diversification,

Table 18
Statistics of k-means clusters for global development indicators of countries.

		Cluster							
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**
	count	4	32	22	9	7	9	7	7
Continent	**unique**	2	5	4	2	3	3	2	2
	top	Asia	America	Africa	Europe	Europe	Asia	Europe	Africa
	freq	3	15	19	7	4	7	5	5
Region	**unique**	4	11	6	5	4	6	5	3
	top	Various regions	Central America	Western Africa	Eastern Europe	Eastern Europe	South-Eastern Asia	Northern/ Southern Europe	Southern Africa
	freq	1	8	10	3	4	3	2	5
Economic situation	**unique**	1	2	2	1	3	2	2	2
	top	High	Upper-middle	Low	High	Upper-middle	Low-middle	High	Upper-middle
	freq	4	18	12	9	5	5	5	4
Mortality rate	**mean**	6.08	7.20	7.45	10.90	11.90	6.57	12.24	9.66
LEB	**mean**	83.22	72.32	63.58	79.29	72.41	72.12	76.56	63.53
Education Exp	**mean**	4.93	4.53	3.65	4.90	4.82	3.85	4.50	7.11
Eletr Access	**mean**	100.00	96.81	50.04	100.00	99.99	99.02	99.96	76.98
Dem	**mean**	74.25	62.91	45.18	90.67	44.00	30.11	85.71	66.71
GDP	**mean**	41735	5181	1204	23498	8393	3491	18046	4479
Inflation	**mean**	0.52	3.99	7.41	1.66	3.42	17.05	2.27	2.54
Unemploy-ment	**mean**	4.14	8.97	4.95	7.60	4.85	7.10	7.86	19.76
FDI	**mean**	14.50	1.69	5.55	1.62	2.13	1.97	2.36	0.93
GHG	**mean**	9.29	3.20	1.55	7.98	11.13	4.75	6.73	4.38
Coping	**mean**	1.75	4.40	6.37	2.30	3.74	4.73	2.90	4.91
CO2	**mean**	0.79	1.61	1.32	1.18	3.90	3.88	0.84	2.09
PM	**mean**	99.91	97.58	99.38	94.29	94.77	99.50	56.02	99.87
Fuel	**mean**	1.680	0.240	0.030	0.790	1.080	0.410	0.370	0.260
Solar	**mean**	0.260	0.020	0.000	0.200	0.040	0.030	0.060	0.030
EPI	**mean**	65.3	42.3	30.3	65.9	48.2	40.8	61.1	39.7

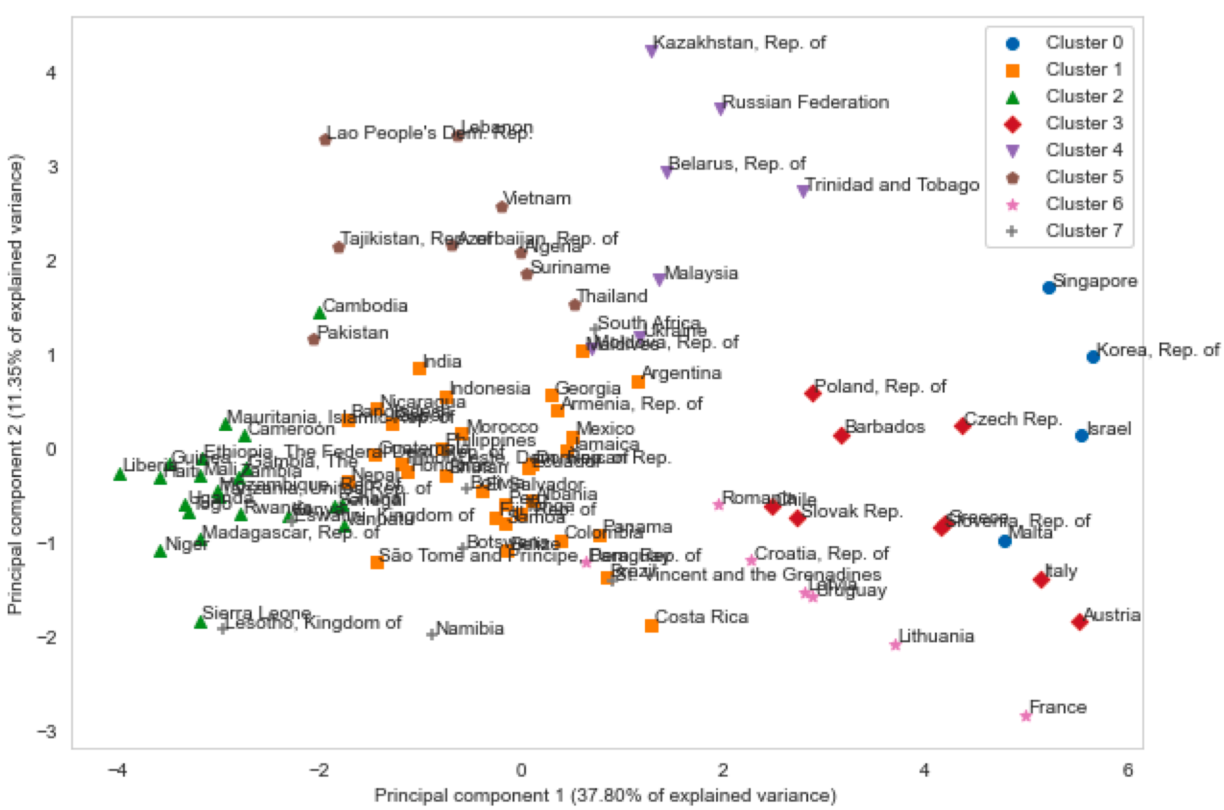

Fig. 5. Cluster graph of countries based on the two principal components of global development indicators.

offers a replicable framework for Kazakhstan’s coal-intensive regions.

d) Cross-Cutting Priorities

Cross-cutting priorities underscore the need for global solidarity and data-driven targeting. Expanding satellite-based air quality monitoring in data-scarce regions (e.g., sub-Saharan Africa, where 80% lack ground stations) would refine policy precision. Meanwhile, advanced economies should fulfill unmet climate finance pledges, such as the COP28 commitment to triple renewables, by redirecting IMF Special Drawing

Rights (SDRs) to resilience projects in low-income clusters. These measures, grounded in empirical clustering insights, offer a roadmap for achieving equitable and sustainable development across diverse national contexts.

5. Conclusions

Adopting a multi-dimensional clustering approach, this study provides a comprehensive analysis of the interrelationships between economic development, environmental quality, and social indicators. By examining a diverse range of variables – spanning economic performance, social conditions, and environmental metrics – our findings offer valuable insights into how these dimensions intersect and vary across different countries.

This study contributes to the literature by (1) integrating economic, social, and environmental indicators into a unified clustering framework, (2) identifying distinct global development patterns that reveal disparities often masked in single-dimension analyses, and (3) proposing a novel 'Global Development and Resilience Index' to benchmark countries' progress.

The results confirm that high-income countries generally exhibit better environmental outcomes and social indicators compared to their lower-income counterparts. Countries in high-income clusters demonstrate robust environmental performance and better social conditions. Conversely, lower-income and emerging economies often face significant challenges related to environmental degradation and social disparities.

The clustering analysis reveals distinct patterns in social indicators. High-income countries are characterised by higher life expectancy, greater access to education, and better overall health outcomes. In contrast, lower-income countries tend to have lower education expenditure, higher mortality rates, and limited access to essential services such as electricity. These disparities underscore the importance of addressing social inequalities as part of sustainable development efforts.

The study highlights substantial differences in environmental performance across countries. High-income nations tend to achieve better environmental outcomes, but often have higher greenhouse gas (GHG) emissions, despite their superior environmental performance indices. However, reliance on fossil fuels remains a common challenge, even among economically advanced countries. On the other hand, many low-income countries, despite having lower overall emissions, struggle with poor air quality and limited capacity to cope with environmental crises. The application of k-means clustering and principal component analysis (PCA) reveals that countries can be grouped into distinct clusters based on their economic, social, and environmental characteristics. These clusters reflect varying levels of development and highlight the complex interplay between economic growth and environmental sustainability. The findings emphasise that a one-size-fits-all approach to policy-making is inadequate, and that instead tailored strategies are needed to address the specific challenges faced by different groups of countries.

Based on the interpretation of the first principal component and its strong correlations with key economic, social, and environmental variables, the resultant scores can be aptly labelled the 'Global Development and Resilience Index'. This index reflects a holistic measure of a country's overall development, social well-being, environmental performance, and resilience to challenges. Countries with the highest PC1 scores, such as the Republic of Korea, Israel, Austria, Italy, and Singapore, are recognised for their robust development and resilience. These nations generally exhibit strong economic performance, high social indicators (such as education and healthcare), as well as effective environmental management, making them well-equipped to handle economic or environmental challenges.

However, it is essential to acknowledge that certain outlier countries which deviate significantly from global averages due to their extreme economic, social, or environmental conditions, were removed from the PCA analysis in order to ensure more accurate and meaningful clustering of the remaining nations. High-income countries such as Ireland, Luxembourg, and Switzerland are economic outliers, due to their exceptionally high GDP per capita, whereas countries such as Venezuela and Zimbabwe are outliers, due to low GDP and extreme inflation. Social outliers such as Serbia and Cuba display unique combinations of high education expenditure, high access to electricity, which contrasts with the other indicators. Similarly, environmental outliers such as Qatar, Bahrain, and Kuwait demonstrate high emissions and poor environmental performance, challenging their overall sustainability.

The insights from this study underscore the need for integrated policy approaches that simultaneously address economic, social, and environmental goals. Policymakers should consider the multifaceted nature of development and tailor interventions to the specific needs and capacities of different countries. For instance, strategies to improve environmental performance in emerging economies might include supporting renewable energy transitions and enhancing coping capacities for environmental challenges. Meanwhile, high-income countries may need to focus on reducing their reliance on fossil fuels and addressing remaining environmental and social inequalities.

In conclusion, achieving sustainable development requires a nuanced understanding of how economic growth, environmental stewardship, and social progress intersect. This study provides a framework for analysing these relationships and offers actionable insights for policymakers whose goal is to foster a more equitable and sustainable global future. As the world continues to confront complex challenges, ongoing research and data-driven approaches will be crucial in guiding effective and informed policy decisions.

While this study offers valuable insights into the multidimensional clustering of countries based on socio-environmental indicators, it is not without limitations. First, the analysis is constrained by the availability and comparability of data across countries, which may lead to the exclusion of some nations or introduce biases. Second, the selected variables, while broad in scope, do not capture all aspects of sustainability, such as biodiversity or cultural factors, which could influence clustering outcomes. Third, the cross-sectional nature of the data also limits the study's ability to account for temporal dynamics. Fourth, the clustering results are sensitive to the choice of variables and normalization techniques; future studies could explore alternative methodologies (e.g., hierarchical clustering) to validate robustness. Finally, while PCA reduces dimensionality, interpreting principal components requires caution, as they may conflate correlated but distinct phenomena. Addressing these limitations in future work could further refine the understanding of global development dynamics.

The identified clusters provide policymakers with a nuanced understanding of how countries compare in terms of environmental, social and economic development. For instance, countries in clusters with low electricity access and education expenditure may benefit from targeted infrastructure investment and capacity building. Similarly, the high performance of countries in more balanced clusters could offer models for integrated development. These findings underscore the importance of context-sensitive policies that align with a country's specific socio-environmental profile rather than one-size-fits-all solutions.

Future studies could build on this work by incorporating dynamic, panel-based datasets to explore how countries transition between clusters over time. Additionally, expanding the indicator set to include more dimensions of sustainability, such as renewable energy adoption, climate vulnerability, or institutional resilience, could enrich the analysis. Finally, applying this clustering approach at sub-national or regional levels may provide more granular insights for localized policymaking.

CRediT authorship contribution statement

Carolina Saraiva: Writing – original draft, Validation, Investigation, Writing – review & editing, Visualization, Methodology, Conceptualization. **Jorge Caiado:** Writing – review & editing, Supervision,

Methodology, Formal analysis, Writing – original draft, Software, Investigation.

Declaration of competing interest

The work described above has not been published previously, it is not under consideration for publication elsewhere, its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

Acknowledgements

Jorge Caiado was supported by Project CEMAPRE/REM - UIDB/05069/2020, financed by FCT/MCTES through national funds.

Appendix 1: Countries retained in principal components analysis (PCA)

Country	Continent	Country	Continent	Country	Continent
Albania	Europe	Guatemala	America	Pakistan	Asia
Algeria	Africa	Guinea	Africa	Panama	America
Argentina	America	Haiti	America	Paraguay	America
Armenia	Asia	Honduras	America	Peru	America
Austria	Europe	India	Asia	Philippines	Asia
Azerbaijan	Asia	Indonesia	Asia	Poland, Rep. of	Europe
Bangladesh	Asia	Israel	Asia	Romania	Europe
Barbados	America	Italy	Europe	Russian Federation	Europe
Belarus	Europe	Jamaica	America	Rwanda	Africa
Belize	America	Kazakhstan, Rep. of	Asia	Samoa	Oceania
Bhutan	Asia	Kenya	Africa	São Tomé and Príncipe	Africa
Bolivia	America	Korea, Rep. of	Asia	Senegal	Africa
Botswana	Africa	Lao	Asia	Sierra Leone	Africa
Brazil	America	Latvia	Europe	Singapore	Asia
Cambodia	Asia	Lebanon	Asia	Slovak Rep.	Europe
Cameroon	Africa	Lesotho, Kingdom of	Africa	Slovenia, Rep. of	Europe
Chile	America	Liberia	Africa	South Africa	Africa
Colombia	America	Lithuania	Europe	St. Vincent and the Grenadines	America
Costa Rica	America	Madagascar, Rep. of	Africa	Suriname	America
Croatia, Rep. of	Europe	Malaysia	Asia	Tajikistan, Rep. of	Asia
Czech Rep.	Europe	Maldives	Asia	Tanzania, United Rep. of	Africa
Dominican Rep.	America	Mali	Africa	Thailand	Asia
Ecuador	America	Malta	Europe	Timor-Leste, Dem. Rep. of	Asia
El Salvador	America	Mauritania	Africa	Togo	Africa
Eswatini	Africa	Mexico	America	Tonga	Oceania
Ethiopia	Africa	Moldova, Rep. of	Europe	Trinidad and Tobago	America
Fiji	Oceania	Morocco	Africa	Uganda	Africa
France	Europe	Mozambique	Africa	Ukraine	Europe
Gabon	Africa	Namibia	Africa	Uruguay	America
Gambia	Africa	Nepal	Asia	Vanuatu	Oceania
Georgia	Asia	Nicaragua	America	Vietnam	Asia
Ghana	Africa	Niger	Africa	Zambia	Africa
Greece	Europe				

Data availability

Data will be made available on request.

References

[1] D.I. Stern, The rise and fall of the environmental Kuznets curve, World Dev. 32 (8) (2004) 1419–1439.
[2] Ö.F. Gültekin, R. Sayar, Y.O. Ari, Socio-economic determinants of environmental degradation: empirical evidence for the environmental Kuznets curve, Econ. Bus. Rev. 9 (3) (2023) 66–97.
[3] A.M. Rasli, M.I. Qureshi, A. Isah-Chikaji, K. Zaman, M. Ahmad, New toxics, race to the bottom and revised environmental Kuznets curve: the case of local and global pollutants, Renew. Sustain. Energy Rev. 81 (2) (2018) 3120–3130.
[4] Horii, R. & Ikefuji, M. (2014) Environment and growth [online]. 37.2014. Fondazione Eni Enrico Mattei (FEEM). Available from: https://www.jstor.org/stable/resrep01049.
[5] M. Onofrei, A.F. Vatamanu, E. Cigu, The relationship between economic growth and CO2 emissions in EU countries: a cointegration analysis, Front. Environ. Sci. (2022) 10, https://doi.org/10.3389/fenvs.2022.934885.
[6] M. Galeotti, Economic growth and the quality of the environment: taking stock, Environ. Dev. Sustain. 9 (2007) 427–454.
[7] B. Almarafi, M. Khudari, A. Abdullah, A critical review of the relationship between environmental performance index, financial development and economic growth, Int. J. Prof. Bus. Rev. 8 (7) (2023) e02568.
[8] T. Khatun, Measuring environmental degradation by using principal component analysis, Environ. Dev. Sustain. 11 (2009) 439–457.
[9] I. Gallego-Alvarez, M.P. Vicente-Galindo, M.P. Galindo-Villardón, M. Rodríguez-Rosa, Environmental performance in countries worlwide: determinant factors and multivariate analysis, Sustainability 6 (11) (2014) 7807–7832.
[10] C. Chen, Q. Ge, Z. Tao, L. Liang, Evaluation of ecological environment quality in Chongqing main city area based on principal component analysis, Sci. Program. (2022) 7374034, https://doi.org/10.1155/2022/7374034.
[11] N. Jain, G. Mohapatra, A comparative assessment of composite environmental sustainability index for emerging economies: a multidimensional approach, Manag. Environ. Qual. 34 (5) (2023) 1314–1331.
[12] R. Kumari, R. Raman, R.K. Patel, Regional disparities in social, environmental, and economic indicators among the Indian sates, GeoJournal 88 (2023) 4351–4371.
[13] Y. Lu, R. Liu, An empirical analysis of Jiangxi Provinces financial development, economic growth, and environmental pollution, in: Wireless Communications and Mobile Computing, 2022, 2022 3327318, https://doi.org/10.1155/2022/3327318.
[14] A. Mathrani, J. Wang, D. Li, X. Zhang, Clustering analysis on sustainable development goal indicators for forty-five Asian countries, Sci 5 (2) (2023) 14, https://doi.org/10.3390/sci5020014.
[15] M. Çağlar, C. Gürler, Sustainable development goals: a cluster analysis of worldwide countries, Environ. Dev. Sustain. 24 (2022) 8593–8624.

[16] M. Jabbari, M.S. Motlagh, K. Ashrafi, G. Abdoli, Differentiating countries based on the sustainable development proximities using the SDG indicators, Environ. Dev. Sustain. 22 (2020) 6405–6423.
[17] Jančovič, P. (2022). Cluster analysis of European Union member states performance in terms of SGD indicators. In: Dančo, J., Kromková, J., Orlovská, S., Pernický, J., Sýkora, J. and Vlková, E., eds. *International Relations 2022: Current issues of world economy and politics*. Bratislava, 1-2 Dezembro 2022. Ekonóm, 251-262.
[18] M. Drastichová, Cluster analysis of sustainable development goal indicators in the European Union, eds., in: M. Bilgin, H. Danis, G. Karabulut, G. Gözgor (Eds.), Euroasian Economic Perspectives: Proceedings of the 25thEuroasia Business and Economics Society Conference, Berlin, Springer, Cham 12, 2020, pp. 99–124.
[19] S. Dasgupta, B. Laplante, H. Wang, D. Wheeler, Confronting the environmental Kuznets curve, J. Econ. Perspect. 16 (1) (2001) 147–168.
[20] W. Antweiler, B.R. Copeland, M.S. Taylor, Is free trade good for the environment? Am. Econ. Rev. 91 (4) (2001) 877–908.
[21] OECD (2012). "Environmental outlook to 2050: the consequences of inaction". Organisation for Economic Co-operation and Development Publishing, https://www.oecd.org/en/publications/oecd-environmental-outlook-to-2050_9789264122246-en.html.
[22] Gough, I. (2017). *Heat, Greed and Human Need: Climate Change, Capitalism and Sustainable Wellbeing*. Edward Elgar Publishing.
[23] D.D. Moran, M. Wackernagel, J.A. Kitzes, S.H. Goldfinger, A. Boutaud, Measuring sustainable development—nation by nation, Ecol. Econ. 64 (3) (2008) 470–474.
[24] Jackson, T. (2009). *Prosperity Without Growth: Economics for a Finite Planet*. Earthscan/Routledge.
[25] Raworth, K. (2017). *Doughnut Economics: Seven Ways to Think Like a 21st-Century Economist*. Chelsea Green Publishing.
[26] World Health Organization (2021) WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide [online]. Geneva: World Health Organization. Available from: https://www.who.int/publications/i/item/9789240034228 [Accessed 1 september 2024].
[27] B. Ratner, The correlation coefficient: its values range between +1/-1, or do they? J. Target. Meas. Anal. Mark. 17 (2009) 139–142.
[28] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (7th ed.). Pearson Education.
[29] Maharaj, E. A., D'Urso, P. & Caiado, J. (2019). Time Series Clustering and Classification, 1 Ed. Boca Raton: CRC Press, Taylor & Francis Group.
[30] I.T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Philos. Trans. R. Soc. A 374 (2065) (2016) 20150202.
[31] H. Fakher, Z. Abedi, Relationship between environmental quality and economic growth in developing countries (based on Environmental Performance Index), Environ. Energy Econ. Res. 1 (3) (2017) 299–310.