Claudia Diana SABĂU-POPA, PhD (corresponding author)
dpopa@uoradea.ro
University of Oradea, Oradea, Romania

Marcel Ioan BOLOȘ, PhD
marcel_bolos@yahoo.com
University of Oradea, Oradea, Romania

Diana Claudia PERȚICAȘ, PhD
dperticas@uoradea.ro
University of Oradea, Oradea, Romania

Victoria BOGDAN, PhD
victoria.bogdan10@gmail.com
University of Oradea, Oradea, Romania

Adrian FLOREA, PhD
adiflorea01@gmail.com
University of Oradea, Oradea, Romania

Mirela BUCUREAN, PhD
mbucurean@gmail.com
University of Oradea, Oradea, Romania

Alexandra-Nicoleta CIUCU (DURNOI), PhD candidate
alexandra.ciucu@csie.ase.ro
Bucharest University of Economic Studies, Bucharest, Romania

A Global Sustainability Development Map. Clustering Countries by Key Indicators

Abstract. *Sustainable development involves a global process aimed at improving economic and environmental conditions in order to increase the quality of human life for current and future generations. The research aims to investigate in the period 2003-2020 a series of indicators intended to define sustainability through its three pillars (environment, economic, social). Using clustering methods, the countries for which data were identified in the studied period were grouped according to their evolution over time with regard to the values recorded to measure sustainability. Additionally, the classes were grouped separately according to the performances recorded in 2020 in order to be able to make a comparison between what happened over time and where they were at the end of the analysed period. Significant differences could be identified between the obtained clusters, especially between areas of Africa and certain states in Asia and the rest of the world. Some of the most impressive and worrying aspects are related to the limited access of African states to electricity, as well as to clean fuels and technologies for cooking. The results showed that 51 countries out of the 80 examined belong to the same classes regardless of the period for which the clustering was chosen, noting that approximately 81.48% of the states that changed*

DOI: 10.24818/18423264/60.1.26.03

their membership classes achieved this migration between the first two groups, which are much more similar compared to the other two clusters.

Keywords: *SDG, cluster analysis, countries, electricity, Ward's method, sustainable development.*

JEL Classification: C38, Q01, W56.

Received: 4 July 2025	Revised: 8 December 2025	Accepted: 2 March 2026

1. Introduction

Sustainable development focuses on a balance between economic growth, social equity, and environmental protection and aims to meet current needs without compromising future generations. Key areas include minimising environmental impact, promoting green technologies, promoting social well-being, and promoting economic justice. Sustainability frameworks, often guided by energy efficiency, resource management, and social development measures, support the long-term resilience of regions. Key principles of sustainable development include (Yue & Hou, 2024) economic sustainability, which involves promoting economic growth that conserves natural resources and protects the environment, improving the quality of life without harming ecosystems social sustainability that ensures fair access to resources and opportunities throughout society and promotes social equality and justice, and ecological sustainability, that emphasises the protection of natural resources and ecosystems and aims to prevent human activities from causing ecological damage (Wu et al., 2022).

The current global economic development is characterised by the importance of digital innovations and sustainable development in all fields. However, the existing literature cannot explain how the environmental, economic and social factors of sustainable development, corresponding to these innovations, affect the new global ecosystem built on the introduction and adaptation of digital technologies and the achievement of the United Nations Sustainable Development Goals (SDGs) at the national level. The article presents the results of evaluating and classifying 80 countries worldwide using 10 SDG indicators to identify clusters and examine important themes. The empirical analysis uses data on a unique combination of indicators, including: clean cooking fuels, electricity access, agricultural land, CO_2 emissions, FDI inflows, GDP growth, inflation, merchandise trade, net migration, and urban population growth. Using cluster analysis implemented via hierarchical algorithms and the Ward method, we analysed the grouping of countries for the year 2020. The presentation of the classes obtained based on these 10 indicators was then followed by the performance recorded by the countries for the period 2003–2020, and finally the differences identified between the four country classes for the two periods mentioned above are presented. We applied the NbClust package and calculated the silhouette index to determine the optimal number of classes.

In the context of globalisation, the risk of uncertainty in both the external and internal environments of a country significantly increases, which can lead to a possible loss of stability. Active exploitation of natural resources, industrial development, and growth in consumption and energy processing lead to a greater environmental impact and contribute to the loss of sustainability. The analysis of the clean cooking fuels indicator shows that fuels derived from plants or animals are expected to replace traditional fuels produced from exhaustible resources with those produced from renewable feedstocks. At the same time, development processes cannot be stopped, and a new development policy must be recognised and defined. National sustainable problems are particularly acute in light of this, which is why I selected the agricultural land indicator and CO_2 emissions. Agriculture is one of the largest sources of greenhouse gas emissions; however, its efficient use could contribute to limiting global warming by the end of the century. OECD experts have evaluated the impact of agriculture on global warming and analysed the policies of developed and developing countries on this issue.

Cluster analysis of the 80 countries identified homogeneous clusters with distinctive systemic characteristics: advanced countries, countries catching up on the gap, stagnating countries, and leaders in sustainable and digital development. This is evidenced by the cross-sectional (2020) and long-term (2003–2020) clustering approach, which highlights both the tandem of structural persistence of the pattern and the transitions between clusters. The identified clusters extend the possibilities for grounding socioeconomic and sustainable strategies at the country-region-industry level, although there are some structural asymmetries. The results showed that 51 countries-maintained membership in the same classes over the analysed period, and significant differences between clusters could be identified, with issues related to electricity, clean fuels, and cooking technologies being the most important.

2. Literature Review

Sustainable development indicators on a macroeconomic level should be analysed through the perspective of exogenous and endogenous influences. Therefore, the present work considered clustering a valuable and appropriate method for analysing and grouping various indicators for achieving sustainable development goals, as well as a valuable tool for comparative analysis of economic and environmental data across countries. Clustering enables researchers to group similar entities based on common characteristics, revealing patterns and insights that would be difficult to detect in ungrouped data. It allows countries with similar profiles to be grouped, facilitating the recognition of shared characteristics and stages of development among groups, thus enabling targeted policy recommendations.

Similar studies recently conducted on clustering countries by achieving SDGs using various indicators were examined and a sample of ten publications was selected to examine the results and main findings (Table 1).

Table 1. Synthesis of selected recent similar studies results

Authors	Title of the paper	Results summary and methodology used
Verma et al. (2020)	A new set of cluster driven composite development indicators	The study demonstrates, using PCA and an innovative application of information filtering and hierarchical clustering, that composite development indicators, constructed from restricted data sets, may omit essential information. Instead, a new set of cluster-based composite development indicators is proposed, capable of enabling valid cross-country comparisons and providing new insights into national development
Wang & Huang (2021)	The impact of COVID-19 pandemic on sustainable development goals – A survey	This research analysed scientific publications from Web of Science related to the COVID-19 pandemic and sustainability, employing bibliometric techniques, cluster analysis, and structural modelling. The results indicate that research in developed countries dominates the field, even though the pandemic affects the sustainable development of developing countries more than that of developed ones. Advanced countries have focused on studying sustainability in education, while developing countries have shown greater attention to economic sustainability during the pandemic.
Magazzino et al. (2022)	Renewable energy consumption, environmental degradation and economic growth: the greener the richer?	This article presents a panel analysis of the relationship between GDP, CO_2 emissions, and the use of renewable energy in the Scandinavian countries over the period 1990–2018. The empirical results suggest that renewable energy consumption is a useful policy instrument for reducing CO_2. The main implication is that CO_2 emissions can be reduced by increasing the use of renewable energy, thereby achieving higher levels of energy efficiency and economic growth.
Wu et al. (2022)	Decoupling of SDGs followed by re-coupling as sustainable development progresses	In this article, the authors employed a correlational-cluster approach and a global database of 166 countries to analyse the evolution of Sustainable Development Goal (SDG) achievement and the interactions among them, measured by the SDG Index. SDG interactions exhibited nonlinear changes as the SDG Index increased: development objectives were both more positively and negatively connected at low and high levels of sustainable development, but clustered into more isolated positive connection clusters at medium levels of sustainable development.
Mathrani et al. (2023)	Clustering Analysis on Sustainable Development Goal Indicators for Forty-Five Asian Countries	The study uses Ward's clustering method to identify commonalities and differences across four key dimensions: economic, social, environmental, and institutional. The results of the study reflect the fact that countries in West Asia had better economic and environmental performance and countries in East and Central Asia had better social and institutional performance than the other regions.
Ahmad & Anwar (2023)	A composite index for sustainable development: measurement and development status of selected countries	Using panel statistical analysis for 140 countries over the period 1995-2020, the authors evaluate and measure the development status of different countries through a composite index for sustainable development, with three sub-indices for economic, social, and environmental dimensions. The study results show that high-income countries achieved stronger economic sustainability scores, while low-income countries exhibited comparatively lower levels of environmental degradation.

Authors	Title of the paper	Results summary and methodology used
Çağlar & Gürler (2022)	Sustainable Development Goals: A cluster analysis of worldwide countries	The study identifies five unique clusters, each reflecting distinct characteristics tied to SDG progress and underlying socioeconomic and cultural conditions. High-income countries generally achieved stronger SDG results, particularly in economic growth, governance, and human development. On the contrary, low-income countries faced challenges in infrastructure and resources, but performed relatively well in environmental areas, such as responsible consumption.
Matenga (2022)	Assessment of energy market's progress towards achieving Sustainable Development Goal 7: A clustering approach	This study analyses global progress toward SDG 7, which seeks to provide affordable, reliable, sustainable, and modern energy for all. The study uses a machine learning approach, employing K-Means clustering to categorise countries based on energy market health indicators such as access, quality, and sustainability.

Source: Authors' processing.

The present work aimed to analyse and group countries according to a set of ten indicators considered to have a considerable impact on the global economy. These indicators include access to clean fuels and cooking technologies, access to electricity, agricultural land, carbon emissions, net inflows of foreign direct investment; annual GDP growth; inflation (measured by the annual GDP deflator), merchandise trade, net migration, and annual urban population growth. In what follows, each indicator of the defined cluster was examined highlighting the relationship between these indicators and the sustainable development goals.

The first indicator, access to clean fuels and cooking technologies, is an essential aspect of sustainable development and has profound implications for public health, environmental sustainability, and social equity. Currently, around 80% of the world's energy needs are met from non-renewable sources. However, increasing energy demand and growing awareness of the risks of climate change have led many countries to call for a transition to sustainable energy, eventually leading to the global community agreeing on the Sustainable Development Goal, which emphasises ensuring access to affordable, reliable, sustainable and modern energy for all (Hollands & Daly, 2023).

Access to electricity is a cornerstone of modern economic development and human well-being, an integral part of Sustainable Development Goal 7 (SDG 7), *"Affordable and clean energy for all"*. This goal set by the United Nations aims to achieve universal access to affordable, reliable, and sustainable energy services by 2030 (Pokhriyal et al., 2022).

The relationship between agricultural land use and the United Nations Sustainable Development Goals is critical to addressing global challenges related to food security, environmental sustainability, and socioeconomic development. Agricultural land is the basis for food production and livelihoods in rural areas and has a direct impact on SDG 2 (Zero Hunger) (Viana et al., 2022) by increasing food availability and supporting economic growth in farming communities.

Foreign direct investment (FDI) plays a complex role in achieving the Sustainable Development Goals in developing regions. Research shows that FDI in areas such as renewable energy and sustainable agriculture supports objectives related to economic growth, clean energy, and food security (Viana et al., 2022).

The connection between economic growth and environmental impact has attracted a great deal of research attention in recent years. While gross domestic product (GDP) growth has traditionally been a primary goal for many countries, this growth often depends on increased resource extraction, energy consumption, and industrial activities, all of which can harm the environment. Adrangi and Kerr's (2022) research on the relationship between GDP growth and the Sustainable Development Goals (SDGs) in emerging economies concludes that a sole focus on GDP may not be sufficient to achieve the SDGs and highlights potential contradictions between economic growth and the Sustainable Development Goals (Brad et al., 2016). Consequently, the relationship between GDP emissions and sustainable development indicators is complex and is influenced by both endogenous and exogenous factors (Sabău-Popa et al., 2024).

Concerning the relationship between net migration and the SDGs, Prada (2020) establishes a clear link between the Sustainable Development Goals (SDGs) and net migration, pointing out that economic indicators such as GDP per capita and unemployment significantly influence migration flows. It highlights the reciprocal relationship in which migration influences the achievement of the SDGs and is simultaneously shaped by the achievement of these goals.

The current paper aims to compare how countries are grouped at the level of a single year (2020) compared to their grouping when considering the evolution over time (the period 2003-2020). This fact presents an element of originality of the paper because most studies focus either on a single year analysis (Çağlar & Gürler, 2022; Kluczek et al., 2025) or on a longer period (Arshad et al., 2020), with a relatively small number of papers being identified that address both methods in a single paper (Chen et al., 2020; Drastichová & Filzmoser, 2019). The reason for choosing cluster analysis refers to identifying the level of homogeneity between countries, it being desirable to determine the similarities between them, as well as the differences, so that comparable policies can be adopted to support sustainability for states that present similar characteristics. In this sense, uniformity is given to countries with a high degree of resemblance, while at the same time taking into account the discrepancies between different clusters, making it easier to notice the strengths or weaknesses of each cluster and act accordingly.

3. Data and methods

The research begins by selecting 10 indicators related to sustainable development applicable at the global level, identifying 80 countries for which data were determined. The research approached the period of interest 2003-2020, this being the common time interval for which values were recorded for all identified countries. Also, the number of indicators was limited to 10 because there was no continuity in data reporting.

Regarding the source of methodological inspiration, this is presented in the form of the study conducted by Wang & Lu (2021) , in which all the steps to be followed to perform a cluster analysis on panel data are explained. Thus, the current research fully follows the steps proposed in the aforementioned paper for the panel clustering.

Cluster analysis involves grouping objects with similar characteristics within classes so that the variance within them is minimal (Mooi & Sarstedt, 2010). For the current research, an unsupervised shape clustering method was used, that of hierarchical algorithms, given the simplicity of the method and its popularity among researchers, with a multitude of studies being identified that affirm its increased use among researchers (Arutyunova & Röglin, 2025; Hexmoor, 2015; Tokuda et al., 2022). In the case of hierarchical methods, they are divided into two categories, agglomerative and divisive. In the case of the first classification, in the initial stage, each object constitutes a class, following which they will join together through successive iterations until they form a single cluster. For divisive methods, the methodology is somewhat reversed, in the sense that it starts from a single cluster that contains all the objects, and then they are divided into classes (Rodriguez et al., 2019). Among the most used binding methods are complete, single, Ward's, and average, each of which has different formulas for grouping objects into classes (Romesburg, 2004).

The first step in the implementation of cluster analysis is data standardisation, this stage being necessary to remove any scale difference problems that may arise between indicators given the fact that there are different units of measurement. Next, Euclidean distances between countries are calculated, this being the most widely used method to calculate distances between objects in clusters (Belyadi & Haghighat, 2021; Grabusts, 2011); moreover, Junthopas and Wongoutong (2025) state that the Euclidean and Manhattan metric methods outperformed Canberra distance. After this stage, two techniques of linking objects in classes were chosen, complete, respectively Ward's method. This choice is justified by the increased performance of these methods over others (Hanadi A. Amhimmid et al., 2025; Vijaya et al., 2019), with most ranking Ward's method as the most efficient, and this linkage method is also considered to be the most frequently used in research (Govender & Sivakumar, 2020). The NbClust package available in RStudio was used to determine the optimal number of classes related to each linkage method applied (Charrad et al., 2014). The next phase consisted of calculating silhouette coefficients to determine the degree of similarity of objects within classes (Zhou & Gao, 2014). To test the efficiency of clustering, the Silhouette method is recognised for its increased frequency of use among researchers (Belyadi & Haghighat, 2021; Gere, 2023). After choosing the linking method and the number of classes, the related dendrogram (CaliŃski, 2014) is illustrated to visualise how the countries are divided into clusters. For easier visualisation of the geographical positioning of the countries in the classes, charts were created with the help of MapChart (MapChart, 2014).

Table 2 shows the 80 countries that were considered in the current study; as can be seen, there are countries from all continents, except Antarctica. At the time of the implementation of this research, 2020 was the last year for which data for the analysed indicators were available. The paper was also limited to 80 countries and 10 indicators because no additional data were identified that coincided with the investigated time period.

Table 2. Country labels

Country	Code	Country	Code	Country	Code	Country	Code
Argentina	ARG	Egypt	EGY	Lithuania	LTU	Poland	POL
Australia	AUS	El Salvador	SLV	Luxembourg	LUX	Portugal	PRT
Austria	AUT	Finland	FIN	Madagascar	MDG	Romania	ROU
Azerbaijan	AZE	France	FRA	Malaysia	MYS	Russian Federation	RUS
Bahrain	BHR	Georgia	GEO	Mauritius	MUS	Saudi Arabia	SAU
Bangladesh	BGD	Germany	DEU	Mexico	MEX	Singapore	SGP
Belarus	BLR	Greece	GRC	Moldova	MDA	Slovenia	SVN
Belgium	BEL	India	IND	Mongolia	MNG	Solomon Islands	SLB
Botswana	BWA	Indonesia	IDN	Morocco	MAR	Spain	ESP
Brazil	BRA	Iraq	IRQ	Mozambique	MOZ	Sri Lanka	LKA
Burkina Faso	BFA	Ireland	IRL	Namibia	NAM	Switzerland	CHE
Cambodia	KHM	Italy	ITA	Nepal	NPL	Tajikistan	TJK
Canada	CAN	Jamaica	JAM	Netherlands	NLD	Tanzania	TZA
Chile	CHL	Japan	JPN	Nigeria	NGA	Thailand	THA
China	CHN	Jordan	JOR	Oman	OMN	Tunisia	TUN
Colombia	COL	Kazakhstan	KAZ	Pakistan	PAK	Ukraine	UKR
Costa Rica	CRI	Kenya	KEN	Panama	PAN	United Kingdom	GBR
Croatia	HRV	Korea, Rep.	KOR	Paraguay	PRY	United States	USA
Czechia	CZE	Kuwait	KWT	Peru	PER	Viet Nam	VNM
Ecuador	ECU	Latvia	LVA	Philippines	PHL	Zimbabwe	ZWE

Source: Authors' processing.

The indicators extracted from the World Bank database for the period 2003-2020 can be seen in Table 3.

Table 3. Indicators

Label	Name	Unit
Ac_ft	Access to clean fuels and technologies for cooking	%
Ac_el	Access to electricity	%
Agr	Agricultural land	%
CO2	CO_2 emissions	metric tons per capita
FDI	Foreign direct investment, net inflows	current US$
GDPg	GDP growth (annual)	%
Inf	Inflation, GDP deflator (annual)	%
Mt	Merchandise trade	% of GDP
N m	Net migration	number
Upg	Urban population growth (annual)	%

Source: Authors' processing using data from World Bank.

Most of the indicators are expressed as a percentage, except for carbon emissions, foreign direct investment, net inflows, and net migration.

4. Results and discussion

The present section presents the results obtained following the application of the clustering methods discussed in the methodology. In the first stage, the grouping of countries for the year 2020 was analysed, then the aim was to present the classes obtained based on the indicators studied through the prism of the performances recorded by the countries for the period 2003-2020. The section culminates with the presentation of the differences identified between the two groups of countries for the two previously mentioned periods.

4.1 Country clustering results for 2020

The first step in the application of cluster analysis was data standardisation, followed by the calculation of Euclidean distances between objects. In the case of hierarchical algorithms, there are several methods of grouping countries into clusters, in the present case the complete and Ward's methods were used. To determine the optimal number of classes related to each method, the NbClust package was applied. Table 4 encapsulated the number of indicators that recommended a certain number of clusters, noting that using the Ward's method, most indicators proposed three classes, while in the case of the complete method, the majority of indicators recommended four clusters.

Table 4. Proposed number of clusters - 2020

Number of clusters	Number of indices (ward)	Number of indices (complete)
2	6	9
3	10	3
4	4	11
5	3	1

Source: Authors' processing.

Looking for the best method of grouping and the optimal number of classes, the silhouette indicator was calculated, a higher value indicating better clustering.

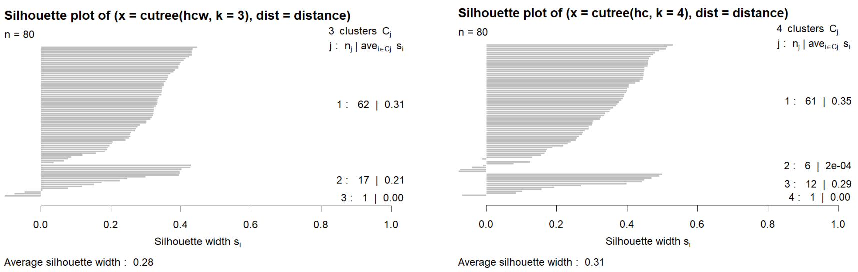

Figure 1. Silhouette - 2020 Ward and complete
Source: Authors' processing using RStudio.

Comparing the two values of the silhouette coefficient presented in Figure 1, it was found that the silhouette coefficient associated with the complete method was higher compared to that of Ward's method, which is why four classes were chosen, as indicated by the complete method.

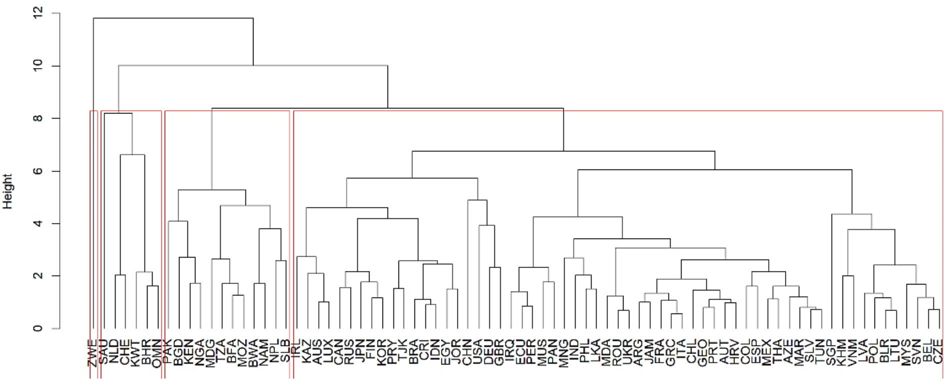

Figure 2. Dendrogram - 2020

Source: Authors' processing using RStudio.

In the dendrogram illustrated in Figure 2, the grouping of countries can be seen, with red highlighting the boundaries between the clusters.

For an easier picture of the classes, the map in Figure 3 was created, to visualise them also from a geographical point of view. It can be seen that most of the countries analysed were grouped into a single class, marked in green. The second class formed, illustrated in blue, was composed of two countries from Europe (the Netherlands and Switzerland) and four other countries from Asia (Bahrain, Kuwait, Oman, and Saudi Arabia). The third class included mostly the countries from Africa (Botswana, Burkina Faso, Kenya, Madagascar, Mozambique, Namibia, Nigeria, and Tanzania) and a few from Asia (Bangladesh, Pakistan, and Nepal) and the Solomon Islands. The fourth class comprised only one country, namely Zimbabwe.

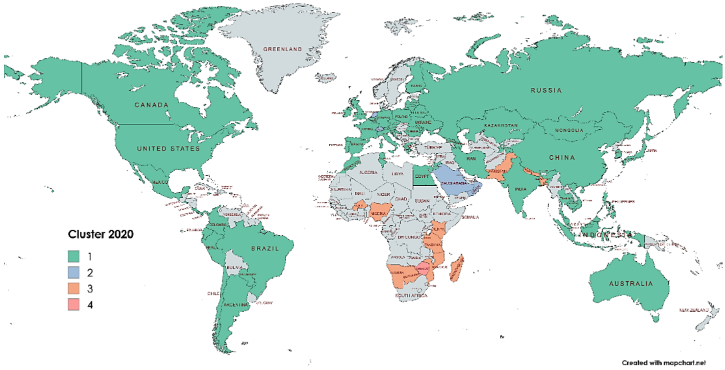

Figure 3. Geographical mapping of countries - 2020

Source: Authors' processing using Mapchart.

According to Table 5, the second class had the highest access to electricity, closely followed by the first class, a sign that most of the analysed states had vast access to this resource. As for the other two classes, much lower values were observed, indicating their limited access to energy. With respect to agricultural area, African countries had the highest values, followed by Zimbabwe, but with a significantly smaller area. The countries with the smallest agricultural area were those that make up the second class, namely the Netherlands, Switzerland, Bahrain, Kuwait, Oman, and Saudi Arabia. As for the characteristics of the countries that form the first class, it can be seen that they had the highest values of foreign direct investment, as well as net migration. Indicators such as carbon emissions, access to clean fuels and technologies for cooking, and merchandise trade, showed lower values compared to the countries in class 2, but much higher than the other two classes. Also, for the GDP growth indicator, only Zimbabwe had a lower value, while for Urban population growth, it has among the lowest values, but higher than the value recorded by the countries of the second class. As far as inflation was concerned, this first grouping was characterised by the lowest value of this indicator compared to the other classes examined.

Table 5. Mean values for groups - 2020

Group	Ac_el	Agr	Ac_ft	CO2	Fdi	GDPg	Inf.	Mt	Nm	Upg
1	0.37	-0.04	0.34	0.00	0.18	-0.10	-0.11	0.06	0.18	-0.24
2	0.40	-0.42	0.65	1.97	-1.49	0.02	-0.24	0.46	-1.35	-1.09
3	-1.89	0.39	-1.92	-0.90	-0.16	0.55	-0.08	-0.48	-0.24	1.72
4	-2.27	0.04	-1.70	-0.93	-0.17	-0.73	8.80	-0.52	-0.23	0.49

Source: Authors' processing.

Class number two presented the highest values of access to clean fuels and technologies for cooking, carbon emissions, and merchandise trade, additionally, it presents the second largest increase in GDP. Moreover, the second cluster displayed the lowest values recorded for foreign direct investment, net inflows, inflation, and net migration, but also for Urban population growth. The third grouping, composed of most of the investigated countries from Africa, along with Bangladesh, Pakistan, Nepal, and the Solomon Islands, was described by the largest increase in GDP as well as in the Urban population. In addition, it registered the second highest value of Foreign direct investment, net inflows, being very close in value to Zimbabwe, but much lower compared to the average of the countries grouped in the first class, which presented the highest value for this indicator. This group also had the second highest inflation, although it was much lower compared to Zimbabwe, which ranks first, and had the second highest value of net migration. Zimbabwe, the country that formed the fourth cluster, was characterised by the lowest values of carbon emissions, merchandise trade, and GDP growth, facing the highest inflation. It had the second highest population growth, with limited access to clean fuels and technologies for cooking, as well as a negative value of net migration, close to that recorded by the countries that composed cluster three.

4.2 Country clustering results for 2003-2020

To carry out this analysis, the steps proposed by Wang & Lu (2021) were followed. As in the previous case, before starting the analysis the data must be standardised. The next step was to identify the degree of factorability, by applying the KMO test. According to this test, the return value was 0.7, indicating an average factorability (Table 6).

Table 6. KMO test

Indicator	**Upg**	**Nm**	**MT**	**Inf.**	**GDPg**	**Fdi**	**CO2**	**Ac_ft**	**Ac_el**	**Agr**	**Overall**
Value	0.76	0.70	0.73	0.87	0.66	0.62	0.77	0.66	0.70	0.64	0.70

Source: Authors' processing.

Subsequent, the common factors were determined by implementing the Principal Component Analysis. In this case, the eigenvalues stand out in the first line, the individual variance of the components, as well as the cumulative variance that they take from the original information (Table 7). To determine the optimal number of components to be taken into the analysis, their cumulative variance must be about 80%, and the eigenvalues have supra-unit values. In this case, to balance both criteria, the first five main components were considered.

Table 7. Principal Component Analysis (PCA)

	PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**	**PC10**
SS loadings	2.98	1.32	1.26	1.01	0.97	0.75	0.65	0.59	0.36	0.12
Proportion Var.	0.30	0.13	0.13	0.10	0.10	0.08	0.07	0.06	0.04	0.01
Cumulative Var.	0.30	0.43	0.56	0.66	0.75	0.83	0.89	0.95	0.99	1.00

Source: Authors' processing.

Consequently, the orthogonal rotation of the previously selected common factors was performed, using the varimax method, to maximise the variance. As can be seen in Table 8, both the eigenvalues and the individual variances had been modified.

Table 8. Varimax factor rotation

	RC1	**RC2**	**RC3**	**RC4**	**RC5**
SS loadings	2.46	1.71	1.27	1.09	1.01
Proportion Var.	0.25	0.17	0.13	0.11	0.10
Cumulative Var.	0.25	0.42	0.54	0.65	0.75

Source: Authors' processing.

With the help of the scores of the common factors and also of the eigenvalues, the comprehensive factor score was calculated, based on which it was possible to implement hierarchical clustering. Following the application of the NbClust method on the comprehensive factor score matrix by implementing the complete linking method, an equal number of indicators proposed 3 and 4 classes, respectively (Table 9).

Table 9. Proposed number of clusters – 2003-2020

Number of clusters	Number of indices (ward)	Number of indices (complete)
2	8	5
3	8	7
4	4	7
5	4	4

Source: Authors' processing.

Ward's binding method was also chosen to determine the best clustering so that the similarity within the classes would be as high as possible. This approach of using hierarchical algorithms by calculating Euclidean distance, and subsequently choosing the Ward method to link objects, has been noted in various studies (Gere, 2023; Teichgraeber & Brandt, 2018) as having a positive impact on the quality of clustering. In this situation, regarding the information presented in Table 9, it was noted that an identical number of indicators choose 2 and 3 classes, respectively.

Figure 4. Silhouette – 3, respectively 4 clusters complete method

Source: Authors' processing.

Since two different clusters were recommended, the silhouette indicators were calculated to determine which clustering had a higher degree of similarity within the clusters. According to Figure 4, it was noted that regardless of whether 3 or 4 classes are chosen, the silhouette coefficient has the same value.

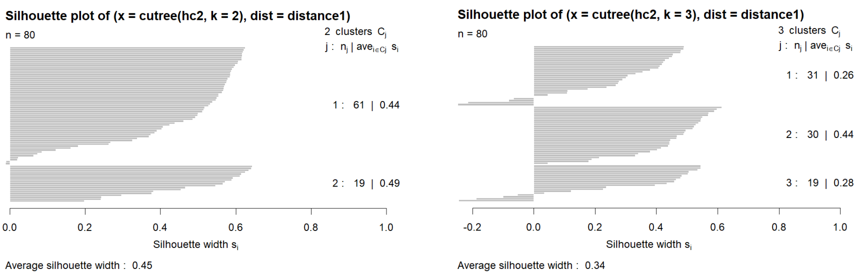

Figure 5. Silhouette – 2, respectively 3 clusters Ward's method

Source: Authors' processing.

As the situation of the groping regarding Ward's method was similar to the one presented for complete linking method, the silhouette indicator for both situations was calculated again, noting from Figure 5 that when countries were grouped into two classes there was a greater similarity of the objects within the clusters.

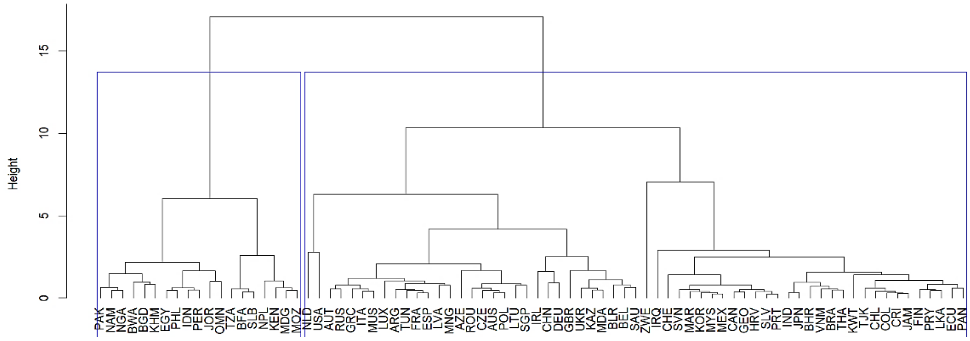

Figure 6. Dendrogram Ward's method

Source: Authors' processing.

In the case of Ward's method, it was observed that grouping into 2 classes led to a lower variance within the classes, noting that 61 states were grouped into one class and 19 into another.

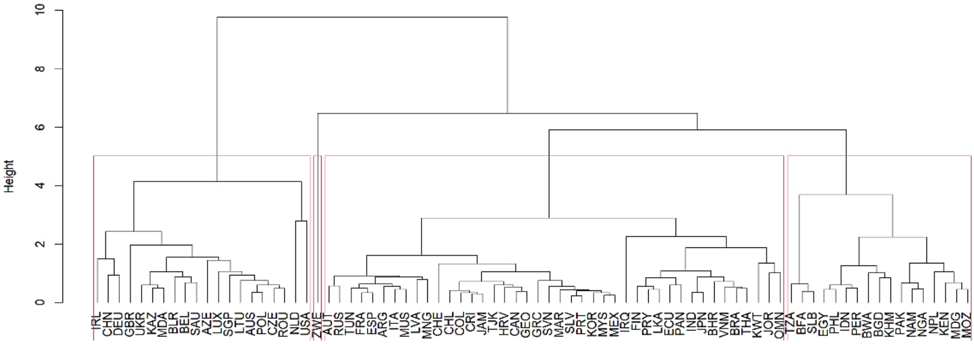

Figure 7. Dendrogram complete method

Source: Authors' processing.

For the complete method, the two grouping methods had equal silhouette coefficients, also, in the case of the clustering carried out for the year 2020, four classes were also considered. Based on these facts, we chose the grouping into four clusters. As in the previous case, Zimbabwe constituted an individual cluster.

Figure 8 illustrated a map of the four clusters identified using the dendrogram in Figure 7 to facilitate visualisation of geographic groupings. Thus, if in the previous situation, in 2020, the clusters somewhat considered a close spatial grouping, in this situation it was noted that the classes no longer take into consideration continental boundaries.

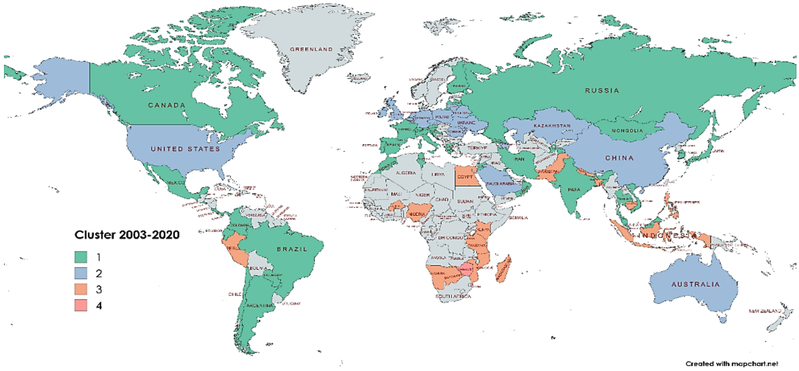

Figure 8. Mapping the clusters for 2003-2020
Source: Authors' processing.

Cluster 2 was composed of countries such as Australia, Azerbaijan, Belgium, Belarus, China, Germany, United Kingdom, Luxembourg, Moldova, Netherlands, Poland, Romania, Singapore, Ukraine and United States, while cluster 3 was made up of Burkina Faso, Egypt, Indonesia, Madagascar, Mozambique, Namibia, Nigeria, Nepal, Pakistan, Peru, Philippines, Solomon Islands and Tanzania. Cluster 1 was represented by 42 countries, which means just over half of the examined states.

Following the calculation of the averages by classes, the values found in Table 10 emerged, according to which cluster 1 was characterised by the second highest value for the indicators, urban population growth, net migration, merchandise trade, foreign direct investment, net inflows, access to clean fuels and technologies for cooking and access to electricity. Except for urban population growth, where class 3 countries had the highest average, for all the other indicators mentioned, cluster 1 was ranked below the average of cluster 2 countries. Also, cluster 1 had the lowest inflation for the analysed period, but also the smallest agricultural area, as well as the second lowest GDP growth per capita. In addition to the five indicators mentioned earlier that cluster 2 had the highest values, it was also observed that the second class has the largest agricultural area. Furthermore, group 2 was characterised by the lowest growth of the urban population, the second lowest inflation, as well as the second highest economic growth per capita. For Cluster 3 it can be said to have had the highest growth in urban population, but also the highest GDP growth per capita. Moreover, it had the second highest inflation, after Zimbabwe, for six other indicators fighting for the lowest values with Zimbabwe. Thus, for merchandise trade, net migration, and access to clean fuels and technologies for cooking, this class was characterised by the lowest values, while for the indicators of foreign direct investment, net inflows, carbon emissions, and access to electricity, only Zimbabwe had lower values compared to those recorded by the average of the countries in this group. Moreover, the group was distinguished by the second smallest agricultural area. For Zimbabwe, the fourth class, it was spotted that it had by far the highest inflation, it is also distinguished by the second largest agricultural area, as well as the second lowest value for indicators such as urban population growth, net migration, merchandise trade, and access to clean fuels and technologies for cooking. Additionally, for the GDP growth, foreign direct investment, carbon emissions, and access to electricity indicators, the country faced the lowest values of all the clusters examined.

Table 10. Mean values for groups, 2003-2020

Group	Upg	Nm	MT	Inf.	GDPg	Fdi	CO2	Ac_ft	Ac_el	Agr
1	-0.14	0.08	-0.07	-0.06	-0.11	-0.16	0.05	0.35	0.39	-0.32
2	-0.53	0.32	0.53	-0.04	0.01	0.61	0.66	0.57	0.50	0.68
3	0.99	-0.55	-0.45	0.03	0.29	-0.30	-0.85	-1.47	-1.44	-0.01
4	-0.43	-0.41	-0.15	3.07	-0.57	-0.35	-0.86	-1.41	-2.08	0.03

Source: Authors' processing.

To determine whether there were significant differences between the two clustering, the one in 2020, and the one for the period 2003-2020, Table 11 was processed, according to which 53 of the 80 states under analysis remained in the same classes regardless of the investigated period. This means that 66.25% of the countries remained in the same cluster. Also, out of the 27 states that changed their class from one analysis to another, 22 of them, i.e., 81.48%, switched between the first two clusters, among these two classes there were smaller differences compared to classes three and four.

Table 11. Cluster modifications

Code	2020	2003-2020	Code	2020	2003-2020	Code	2020	2003-2020	Code	2020	2003-2020
ARG	1	1	EGY	1	3	KWT	2	1	PER	1	3
AUS	1	2	ESP	1	1	LKA	1	1	PHL	1	3
AUT	1	1	FIN	1	1	LTU	1	2	POL	1	2
AZE	1	2	FRA	1	1	LUX	1	2	PRT	1	1
BEL	1	2	GBR	1	2	LVA	1	1	PRY	1	1
BFA	3	3	GEO	1	1	MAR	1	1	ROU	1	2
BGD	3	3	GRC	1	1	MDA	1	2	RUS	1	1
BHR	2	1	HRV	1	1	MDG	3	3	SAU	2	2
BLR	1	2	IDN	1	3	MEX	1	1	SGP	1	2
BRA	1	1	IND	1	1	MNG	1	1	SLB	3	3
BWA	3	3	IRL	1	2	MOZ	3	3	SLV	1	1
CAN	1	1	IRQ	1	1	MUS	1	1	SVN	1	1
CHE	2	1	ITA	1	1	MYS	1	1	THA	1	1
CHL	1	1	JAM	1	1	NAM	3	3	TJK	1	1
CHN	1	2	JOR	1	1	NGA	3	3	TUN	1	1
COL	1	1	JPN	1	1	NLD	2	2	TZA	3	3
CRI	1	1	KAZ	1	2	NPL	3	3	UKR	1	2
CZE	1	2	KEN	3	3	OMN	2	1	USA	1	2
DEU	1	2	KHM	1	3	PAK	3	3	VNM	1	1
ECU	1	1	KOR	1	1	PAN	1	1	ZWE	4	4

Source: Authors' processing.

In the remaining five cases, Egypt, India, Cambodia, Peru, and the Philippines were in class 1 for 2020 and moved to class 3 for 2003-2020.

This article examines the issues of cluster development, namely the key factors that ensure this process, a systematic study of the specialised literature and reports, an analysis of cluster activities, as well as the results that support the SDG monitoring framework. In this sense, we can highlight the energy access deficit in African and South Asian countries, as well as Zimbabwe's extreme vulnerability to inflation. The analysis contributes to outlining the determinants of divergence, including migratory flows and the dynamics of SDIs, from the perspective of environmental pressure.

The practical basis of the study lies in the following grouping of elements, such as access to energy, whose constraints limit the accumulation of human capital and productivity, and which comprehensively reveal all aspects of the sustainable development of the global economy based on migration. Positive aspects may signal economic attractiveness or demographic advantages and may explain the link between the dynamics of foreign investment flows, technological progress, and institutional quality.

5. Conclusions, limits, and outlooks

The present study focused on clustering 80 worldwide countries upon achieving SDGs based on relevant macroeconomic indicators. During the research, significant differences could be identified between the obtained clusters, especially between areas of Africa and certain states in Asia and the rest of the world. Some of the most impressive and worrying aspects are related to the limited access of African states to electricity, as well as to clean fuels and technologies for cooking. Also, the alarming inflation in Zimbabwe is another very worrying aspect, considering that not only for the year 2020 it shows a huge value, but also for the period 2003-2020. Another aspect worth mentioning is the fact that 51 countries out of the 80 examined belong to the same classes regardless of the period for which the clustering was chosen, noting that approximately 81.48% of the states that changed their membership classes achieved this migration between the first two groups, which are much more similar compared to the other two clusters obtained.

Compared to similar studies that used the clustering method (Mathrani et al., 2023; Hanadi et al., 2025; Chen et. al., 2020), the present research focused on investigation of 80 worldwide countries using 10 SDG indicators to identify clusters, to examine the salient themes and to search differences between the four classes of countries for the year 2020, respectively, 2003-2020, which also brings novelty and originality to our research. Using cluster analysis implemented through hierarchical algorithms and Ward's method, we analysed the grouping of countries for the year 2020, then the presentation of the classes obtained based on the 10 indicators studied was followed through the performance recorded by the countries for the period 2003-2020. We applied the NbClust package and calculated the silhouette indicator was calculated, in order to determine the optimal number of classes. Findings proved that 51 countries preserved their membership to the same classes for the analysed time and significant differences could be identified between clusters, issues on electricity, clean fuels, and cooking technologies were the more salient.

If Cluster 1 that contained the majority of the countries analysed in both periods, in 2020 had the highest values of foreign direct investment, as well as net migration and the lowest values of inflation, in 2003-2020 is surpassed by cluster 2 that has the highest value for the indicators, urban population growth, net migration, merchandise trade, foreign direct investment, net inflows, access to clean fuels and technologies for cooking and access to electricity.

In both periods, the cluster 2 is characterised by the lowest growth of the urban population, the largest agricultural area, the highest access to electricity, the second lowest inflation, as well as the second highest economic growth per capita. The third grouping, composed of most of the countries analysed in Africa, is described by the highest growth of GDP as well as in the Urban population. Zimbabwe, the country that forms the fourth cluster in both periods, is characterised by the lowest values of carbon emissions, merchandise trade, and GDP growth, facing the highest inflation.

For the grouping of countries into clusters according to the evolution in the period 2003-2020, a significant difference is noted between the number of elements in the clusters, the first containing more than half of them and the second another 15 clusters. The first two clusters had much more similar results compared to the other two, the differences between them and the others being very pronounced for most indicators, which is why this method of grouping is not surprising. It can be seen that in the case of 2020 these are more unevenly distributed, most of them being classified in cluster 1 (61), and another six in cluster 2, the first two clusters having much more similar characteristics compared to the other two clusters. This framing can be explained by the effects of policies and initiatives related to sustainability that have already been implemented with the aim of helping the sustainable development of countries, providing greater uniformity of characteristics, this element being specific to convergence (Borowiec & Papież, 2024), in which countries with poorer results experience more spectacular growth over time, being helped to reach the level of better performing countries.

The novelty and originality of our study lies in the selection and use of the 10 indicators through which the level of sustainable development in the 80 countries analysed is assessed: access to clean fuels and technologies for cooking, access to electricity, agricultural land, CO2 emissions, Foreign direct investment, GDP growth, Inflation, Merchandise trade, Net migration, Urban population growth, a unique combination that has not been used in other similar studies. Another element of novelty and originality lies in the evaluation of the grouping of the 80 countries through clusterisation for the period 2003-2020, compared with the clustering according to the values of the 10 indicators for the year 2020.

The results of this study are useful in formulating public policies, appropriate for reducing CO_2 emissions and ensuring sustainable development of economies both nationally and globally.

The limitations of the study are addressed to a relatively limited number of countries to be able to discuss the world evolution regarding the analysed indicators, but it is worth noting that, at the time of the research, no data related to the other countries of the world were identified for all the indicators and the examined period. Also, the choice of the clustering method brings with it changes in the structure of the classes, in the sense that by implementing another method the results obtained may differ. However, ameliorating this limitation is the fact that, as previously mentioned, more than 66% of the countries were affiliated with the same classes regardless of the investigated period.

References

[1] Adrangi, B., Kerr, L. (2022), *Sustainable Development Indicators and Their Relationship to GDP: Evidence from Emerging Economies. Sustainability*, 14(2), 658, https://doi.org/10.3390/su14020658.

[2] Ahmad, A., Anwar, S. (2023), *A Composite Index for Sustainable Development: Measurement and Development Status of Selected Countries. Journal of Economic Impact*, 5(1), 1-14, https://doi.org/10.52223/jei5012301.

[3] Arshad, Z., Robaina, M., Botelho, A. (2020), *The role of ICT in energy consumption and environment: An empirical investigation of Asian economies with cluster analysis. Environmental Science and Pollution Research*, 27(26), 32913-32932, https://doi.org/10.1007/s11356-020-09229-7.

[4] Arutyunova, A., Röglin, H. (2025), *The Price of Hierarchical Clustering. Algorithmica*, 87(10), 1420–1452, https://doi.org/10.1007/s00453-025-01327-7.

[5] Belyadi, H., Haghighat, A. (2021), *Unsupervised machine learning: Clustering algorithms*. In Machine Learning Guide for Oil and Gas Using Python (pp. 125-168). Elsevier, Amsterdam, Netherlands, https://doi.org/10.1016/B978-0-12-821929-4.00002-0.

[6] Borowiec, J., Papież, M. (2024), *Convergence of CO2 emissions in countries at different stages of development. Do globalisation and environmental policies matter? Energy Policy*, 184, 113866, https://doi.org/10.1016/j.enpol.2023.113866.

[7] Brad, S., Mocan, B., Brad, E., Fulea, M. (2016), *Environmentally Sustainable Economic Growth. Amfiteatru Economic*, 18(42), 446-460.

[8] Çağlar, M., Gürler, C. (2022), *Sustainable Development Goals: A cluster analysis of worldwide countries. Environment, Development and Sustainability*, 24(6), 8593-8624, https://doi.org/10.1007/s10668-021-01801-6.

[9] CaliŃski, T. (2014), *Dendrogram*. In R.S. Kenett, N.T. Longford, W.W. Piegorsch, F. Ruggeri (Eds), Wiley StatsRef: Statistics Reference Online (1st edn). Wiley, Hoboken, New Jersey, USA, https://doi.org/10.1002/9781118445112.stat05624.

[10] Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A. (2014), *NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. Journal of Statistical Software*, 61(6), https://doi.org/10.18637/jss.v061.i06.

[11] Chen, H., Zhang, L., Zou, W., Gao, Q., Zhao, H. (2020), *Regional differences of air pollution in China: Comparison of clustering analysis and systematic clustering methods of panel data based on gray relational analysis. Air Quality, Atmosphere & Health*, 13(10), 1257-1269, https://doi.org/10.1007/s11869-020-00880-0.

[12] Drastichová, M., Filzmoser, P. (2019), *Assessment of sustainable development using cluster analysis and principal component analysis. Problemy Ekorozwoju*, 14(2), 7-24.

[13] Gere, A. (2023), *Recommendations for validating hierarchical clustering in consumer sensory projects. Current Research in Food Science*, 6, 100522, https://doi.org/10.1016/j.crfs.2023.100522.

[14] Govender, P., Sivakumar, V. (2020), *Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). Atmospheric Pollution Research*, 11(1), 40–56, https://doi.org/10.1016/j.apr.2019.09.009.

[15] Grabusts, P. (2011), *The Choice of Metrics for Clustering Algorithms*. *Proceedings of the 8th International Scientific and Practical Conference*, Rēzekne, Latvia, June 2011, 11, 70-76.

[16] Hanadi A. Amhimmid, Fatma Alzahra A. Aljehany, Mohamed A. Mohamed, Kasem A. Farag. (2025), *Comparative Study of Four Methods in Hierarchical Cluster Analysis. Libyan Journal of Medical and Applied Sciences*, 11-16, https://doi.org/10.64943/ljmas.v3i4.183.

[17] Hexmoor, H. (2015), *Diffusion and Contagion. In Computational Network Science* (pp. 45–64). Elsevier, Amsterdam, Netherlands, https://doi.org/10.1016/B978-0-12-800891-1.00006-8.

[18] Hollands, A.F., Daly, H. (2023), *Modelling the integrated achievement of clean cooking access and climate mitigation goals: An energy systems optimization approach. Renewable and Sustainable Energy Reviews*, 173, 113054, https://doi.org/10.1016/j.rser.2022.113054.

[19] Junthopas, W., Wongoutong, C. (2025), *Pre-Determining the Optimal Number of Clusters for k-Means Clustering Using the Parameters Package in R and Distance Metrics. Applied Sciences*, 15(21), 11372, https://doi.org/10.3390/app152111372.

[20] Kluczek, A., Woźniak, A., Żegleń, P. (2025), *Uncovering economic impacts and dynamics of European energy policy: Evidence from DEMATEL, panel data, and cluster analysis. PLoS One*, 20(11), e0322525.

[21] Magazzino, C., Toma, P., Fusco, G., Valente, D., Petrosillo, I. (2022), *Renewable energy consumption, environmental degradation and economic growth: The greener the richer? Ecological Indicators*, 139, 108912, https://doi.org/10.1016/j.ecolind.2022.108912.

[22] MapChart. (2014), *MapChart*, https://www.mapchart.net/about.html.

[23] Matenga, Z. (2022), *Assessment of energy market's progress towards achieving Sustainable Development Goal 7: A clustering approach. Sustainable Energy Technologies and Assessments*, 52, 102224, https://doi.org/10.1016/j.seta.2022.102224.

[24] Mathrani, A., Wang, J., Li, D., Zhang, X. (2023), *Clustering Analysis on Sustainable Development Goal Indicators for Forty-Five Asian Countries. Sci*, 5(2), 14, https://doi.org/10.3390/sci5020014.

[25] Mooi, E., Sarstedt, M. (2010), *Cluster Analysis*. In E. Mooi & M. Sarstedt, A Concise Guide to Market Research (pp. 237-284), Springer Berlin Heidelberg, Germany, https://doi.org/10.1007/978-3-642-12541-6_9.

[26] Pokhriyal, N., Letouzé, E., Vosoughi, S. (2022), *Accurate intercensal estimates of energy access to track Sustainable Development Goal 7. EPJ Data Science*, 11(1), 60, https://doi.org/10.1140/epjds/s13688-022-00371-5.

[27] Prada, E.-M. (2020), *The Relationship Between Sustainable Development Goals and Migration. An EU-28 Perspective. Journal of Social and Economic Statistics*, 9(1), 28-45, https://doi.org/10.2478/jses-2020-0004.

[28] Rodriguez, M.Z., Comin, C.H., Casanova, D., Bruno, O.M., Amancio, D.R., Costa, L.D.F., Rodrigues, F.A. (2019), *Clustering algorithms: A comparative approach. PLOS ONE*, 14(1), e0210236. https://doi.org/10.1371/journal.pone.0210236.

[29] Romesburg, H.C. (2004), *Cluster Analysis for Researchers*. Lulu Press, Morrisville, USA.

[30] Sabău-Popa, C.D., Bele, A.M., Bucurean, M., Mociar-Coroiu, S.I., Tarcă, N.N. (2024), *A Panel Analysis Regarding the Influence of Sustainable Development Indicators on Green Taxes. Sustainability*, 16(10), 4072. https://doi.org/10.3390/su16104072.

[31] Teichgraeber, H., Brandt, A.R. (2018), *Systematic Comparison of Aggregation Methods for Input Data Time Series Aggregation of Energy Systems Optimization Problems*. In Computer Aided Chemical Engineering, 44, 955–960, Elsevier, Amsterdam, Netherlands, https://doi.org/10.1016/B978-0-444-64241-7.50154-3.

[32] Tokuda, E.K., Comin, C.H., Costa, L.D.F. (2022), *Revisiting agglomerative clustering. Physica A: Statistical Mechanics and Its Applications*, 585, 126433, https://doi.org/10.1016/j.physa.2021.126433.

[33] Verma, A., Angelini, O., Di Matteo, T. (2020), *A new set of cluster driven composite development indicators. EPJ Data Science*, 9(1), 8, https://doi.org/10.1140/epjds/s13688-020-00225-y.

[34] Viana, C.M., Freire, D., Abrantes, P., Rocha, J., Pereira, P. (2022), *Agricultural land systems importance for supporting food security and sustainable development goals: A systematic review. Science of The Total Environment*, 806, 150718, https://doi.org/10.1016/j.scitotenv.2021.150718.

[35] Vijaya, Sharma, S., Batra, N. (2019), *Comparative Study of Single Linkage, Complete Linkage, and Ward Method of Agglomerative Clustering*. International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), February 2019, Faridabad, India, 568–573, https://doi.org/10.1109/COMITCon.2019.8862232.

[36] Wang, Q., Huang, R. (2021), *The impact of COVID-19 pandemic on sustainable development goals – A survey. Environmental Researc*h, 202, 111637, https://doi.org/10.1016/j.envres.2021.111637.

[37] Wang, W., Lu, Y. (2021), *Application of Clustering Analysis of Panel Data in Economic and Social Research Based on R Software. Academic Journal of Business & Management*, 3(10), https://doi.org/10.25236/AJBM.2021.031018.

[38] Wu, X., Fu, B., Wang, S., Song, S., Li, Y., Xu, Z., Wei, Y., Liu, J. (2022), *Decoupling of SDGs followed by re-coupling as sustainable development progresses. Nature Sustainability*, 5(5), 452-459, https://doi.org/10.1038/s41893-022-00868-x.

[39] Yue, N., Hou, J. (2024), *A Study of GGDP Transition Impact on the Sustainable Development by Mathematical Modelling Investigation. Mathematics*, 12(19), 3005, https://doi.org/10.3390/math12193005.

[40] Zhou, H.B., Gao, J.T. (2014), *Automatic Method for Determining Cluster Number Based on Silhouette Coefficient. Advanced Materials Research*, 951, 227-230, https://doi.org/10.4028/www.scientific.net/AMR.951.227.