SEEDS
Sustainability Environmental Economics and Dynamics Studies

# Working Paper Series

*Clustering environmental performances, energy efficiency and clean energy patterns: a comparative static approach across EU Countries*

by

Marco Quatrosi

07/2022

SEEDS is an interuniversity research centre. It develops research and higher education projects in the fields of ecological and environmental economics, with a special focus on the role of policy and innovation. Main fields of action are environmental policy, economics of innovation, energy economics and policy, economic evaluation by stated preference techniques, waste management and policy, climate change and development.

The SEEDS Working Paper Series are indexed in RePEc and Google Scholar.
Papers can be downloaded free of charge from the following websites:
http://www.sustainability-seeds.org/.

Enquiries:info@sustainability-seeds.org

SEEDS Working Paper 07/2022
July 2022
By Marco Quatrosi

# Clustering environmental performances, energy efficiency and clean energy patterns: a comparative static approach across EU Countries

Marco Quatrosi (qtrmrc@unife.it)

Department of Economics and Management, University of Ferrara, Via Voltapaletto, 11 ,44121 Ferrara (Emilia-Romagna, Italy)

**Abstract**

In the context of convergence of objectives among the single Member States within the European Union, environmental policy has always been considered one pivotal and necessary step towards a cohesive EU. Employing clustering techniques, this work identifies affinities in environmental performances (e.g., $CO_2$ emissions), energy efficiency, and clean energy patterns for European countries. K-medoids clustering will be used for a cross-section of the total carbon dioxide emission in three reference years (2008, 2013, 2018). Data to feed the algorithm have been selected considering the well-established IPAT relationship as an analytical framework. After preliminary analysis, results highlighted the presence of persistent groups of countries over time with marked characteristics in terms of environmental performances, energy efficiency, and clean energy patterns. Considering the limitations of data employed and the potentialities of the methodological approach, this work could shed light on a new perspective of analysis in light of the harmonization path the EU has been undertaking since its foundation. These findings could better address policymakers in terms of convergence of environmental policy implementing new measures to promote low-carbon consumption and production patterns with a specific focus on energy efficiency (e.g., heating and cooling) and sustainable sources (e.g., nuclear power).

**Keywords:** IPAT; clustering; renewables; energy efficiency; emissions; environmental policy; energy policy; energy mix

**JEL Codes:** Q50; Q43; C38

## 1.1 Introduction

The process of sustainability transition has always witnessed Europe as one of the most ambitious players in the global landscape. In this sense, while finding an alignment at the international level, the EU is also seeking out harmonization of environmental policy objectives at the Member States' level. This process will strengthen the EU's position as an international leader, promoting a more conscientious relationship between humans and the environment. The European Union has been committed to reducing GHGs emissions (EU2020 Climate and Energy Package). However, only 13%[1] of the overall energy supply in the Union comes from renewable sources. Besides the idea of a worldwide carbon tax, Europe has put a cap on its level of emissions (e.g., European Emission Trading Scheme), with the power sector being among the most prominent players. The power sector is also transitioning to more efficient use of energy along with more sustainable energy sources. Estimates in DG Energy (2020) show that in Europe, in 2008, renewable generation has outperformed fossil fuel-based energy production.

The IPAT relationship represents a simple yet effective framework of analysis bridging human beings and nature. The underlying philosophy of the relationship, as stated by Ehrlich and Holdren (1971) in their seminal work, aims to assess the (negative) impact society is exerting on the environment. Empirically, the IPAT model has considered energy about the technology domain save more recent specifications. With some exceptions, the discourse around the energy transition investigates the influence of its main dimensions (e.g., energy efficiency, renewable sources) concerning the other elements of the IPAT relation (Chontanawat, 2018). According to the diverse epistemological viewpoint collected by Chertow (2000), this influence could be assessed via the contribution either incrementing environmental pollution or reducing. From a strictly applied point of view, IPAT has been employed to assess which dimension exerts the higher contribution on the environmental impact, implying a causal relationship. However, the IPAT is also useful to decompose those effects over time and space, providing a synthetic measure of the main drivers of (human) impact on the environment. In this sense, a wide range of empirical applications have been employed at different territorial levels. However, this dimension could be further explored to highlight common patterns of sustainability transition. In this sense, clustering techniques have been proved to identify hidden patterns of data homogeneity without pre-existing knowledge of the relationship.

While other algorithms somewhat imply an underlying relationship among data (supervised learning), clustering is instead applied to uncover relationships (unsupervised learning). Unlike other statistical techniques, clustering thus solves the issue of dimensionality reduction without significant loss of information (Kaufman and Rousseeuw, 2005). Clustering algorithms can be employed in several fields to find homogeneity in data with many diverse characteristics. In the specific context, this work will apply a clustering algorithm to find common patterns on a set of dimensions related to IPAT with a focus on energy. The data employed come from 31 European Countries (27 EU + Norway, Lichtenstein, Iceland, United Kingdom) spanning from economic (GDP per capita), industrial (% of the manufacturing sector in VA in GDP) demographic (population, density, urbanization), environmental ($CO_2$ emissions from fossil fuel combustion), energy (primary energy consumption, renewable energies, share of fossil fuels in electricity production). A clustering algorithm will be performed on a cross-section of this dataset for three reference periods (2008, 2013, 2018). Using a comparative static approach, different time frames will monitor clusters' (possible) evolution and their characteristics over time. The context of the EU provides valuable insights as it will be possible to assess whether the EU-wide policy framework has somehow affected the harmonization of the existing differences among the Member States. The environmentalist-industrial ecologists' debate will provide ulterior insights towards a clearer understanding of the IPAT relationship. The simplicity yet effectiveness of IPAT makes it a potential tool for policymakers tackling a complex issue such as the human-environment system.

The paper will review the literature on applications of the IPAT relationship and clustering techniques. More details will then be provided regarding the specific clustering algorithm and the data employed. Cluster analysis identified homogeneous countries with diverse environmental, economic, and demographic characteristics. On the other hand, the comparative static analysis of the three periods identified some groups

[1] Source Eurostat, data from 2017

of countries within clusters that have been stable over time. Northern European countries have scored better environmental performances in emissions and clean energy consumption. The same stands for the Baltic States, even though they reached that cluster only in 2018. Despite the relative homogenization over the decades, a well-established cluster (e.g., Germany, France, Italy, Spain, United Kingdom) has emerged, showing relatively worse environmental performances with a higher impact on the population dimension. Belgium, Luxembourg, and the Netherlands have a relatively high wealth per capita and room for improvement in reducing environmental pressures. For the most, Balkan and Eastern European States score a relatively poor amount of wealth and a promising level of clean energy consumption and energy efficiency. However, the cluster scored the highest share of electricity production from (liquid and solid) fossil fuels. In fact, among those countries, only Croatia managed to move to the cluster with higher environmental performances in 2018. The transition process over the decade for those countries was identified in 2008 when Croatia, Greece, Ireland, and Portugal formed the fifth cluster. The case of Poland could be considered an opposite pattern, as the Country returned among Balkan and Eastern European States in 2018 after a transition phase in 2013. More substantial discrimination among countries could be identified in the demographic dimension, with Germany, France, Italy, and the United Kingdom presenting a relatively higher population level. Overall, the efforts at the policy level to harmonize the Member States from a policy perspective have produced certain results for specific cases. However, cluster analysis is highly dependent on the choice of data employed to feed the algorithms. The IPAT relationship has been chosen to provide a well-established analytical framework for variable selection. Further expansions of the analysis will include variables more directly related to the objective set at the EU level on low-carbon transition.

## 2.1The IPAT relationship

The relationship between economic growth and environmental pressure resulting in global warming and a potential threat to human life and livelihood has been consolidated on the scientific ground. Historically, the scientific community has been striving to propose several frameworks to encompass human and environmental systems. Over the decades, McNicoll (2015) notices the IPAT relationship has been further refined considering the impact I as influenced by population P, affluence A, and technology T. More accurate formalizations also led to different interpretations of IPAT from the one proposed by Ehrlich and Holdren (1971) in their seminal paper. While environmentalists think population growth is at the core of environmental pressure, industrial ecologists think income increases improve quality of life (da Silva et al., 2019). In this sense, some applicative instances of the IPAT model include some measure of energy efficiency and energy transition as a potential factor influencing human impact on the environment (IPAT-E). Brizga et al. (2013) used energy intensity and total primary energy supply to investigate carbon dioxide drivers of the former Soviet Union. Yue et al. (2013), in the context of IPAT, extended the model including *emergy* analysis to assess the sustainability of a Chinese industrial province. On the other hand, Chontanawat (2018) employed energy efficiency measures in a decomposition analysis of $CO_2$ emissions for ASEAN countries. Wen and Li (2019) introduce several dimensions of energy consumption in a structural equation model assessing potential driving forces for carbon dioxide emission at China’s regional and national levels.

## 3.2 Clustering

Clustering algorithms have been either used to find structures among data or to synthetize information more neatly. In this sense, clustering algorithms have been widely applied to find common patterns in data (Kaufman and Rousseeuw 2005). Clustering algorithms have mainly been applied to combine economic and non-economic objects of different orientations, (Babenko et al. 2021; Banga and Sinha 2018; Boumans and Leonelli 2020; Dogan and Birant 2021; Jayatilake and Ganegoda 2021). Algorithms can aggregate computing the distance among the diverse characteristics of data. The different algorithm classes partition data according to a specific distance measure (e.g., Euclidean, Manhattan) Steinley (2006). Data can be clustered according to a decision-tree structure (e.g., *hierarchical clustering*) or based on local proximity (e.g., *density-based clustering*). Among density-based algorithms, k-means, one of the widely used attempts to minimize the squared distance of some *k* objects, yielding to centroids. On the other hand, k-medoids selects k objects around which building the clusters in either small (*Partitioning Around Medoids*) or large (*Clustering Large Applications*) datasets. Clustering can be performed matching a (statistical) model-based structure (e.g., *model-*

*based clustering*). Homogeneous data aggregation can also come out of graph representations and computations (e.g., *graph-based clustering*). In this sense, clustering can be employed to enhance similarities among data or to mark dissimilarities. In environmental analysis, Franceschi et al. (2018) employed clustering algorithms to analyze the concentration of PM10 and PM2.5. K-means algorithm was used to cluster different simulations to assess the risk of atmospheric pollution (Cervone et al. 2008). On the other hand, Di et al. (2019) used partitioning around medoids (PAM) and expected-maximization (E-M) clustering algorithm in analysing wastewater heavy pollutant contents. Chang and Lee (2019) applied the k-means algorithm to compute the Sustainable Development Progress Index (SDPI) with data of 32 OECD Countries. A combination of k-means and kernel-based (density-based spatial clustering analysis with noise- DBSCAN) has been employed to analyse the supply and demand flow of dockless shared bicycles Chen and Chen (2020).

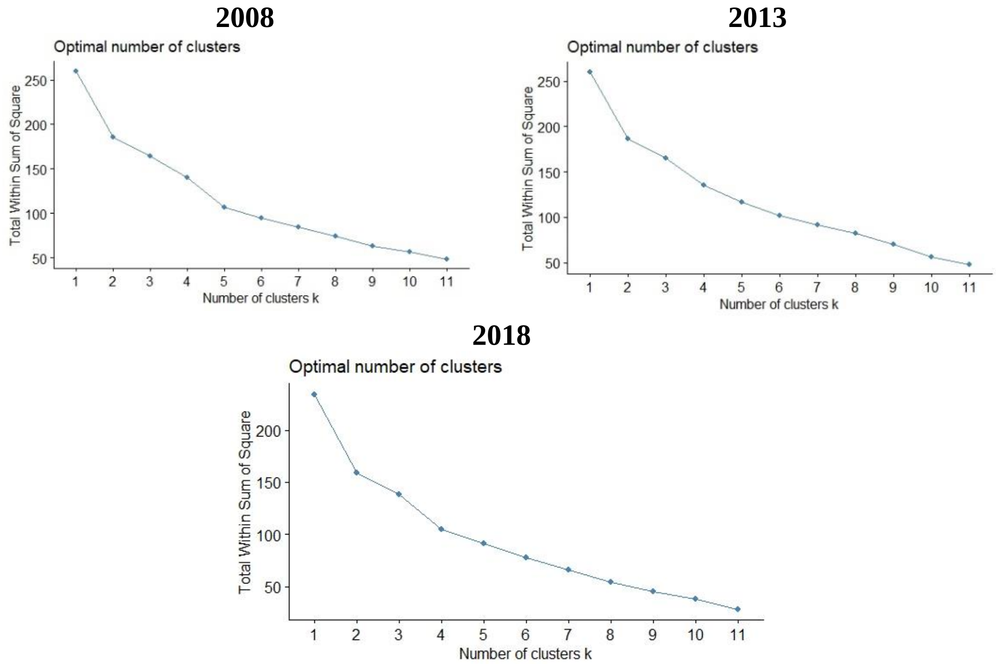
## 4.1 Data and Methodology

The choice of the data to perform the analysis follows the one traditionally employed in the literature for the IPAT identity. For a proxy of environmental pressure, $CO_2$ emission series have been calculated from the combustion of fossil fuels, following the methodology in Eggleston et al. (2006); Quatrosi (2020). The choice of emissions from fuel combustion will provide a more direct link with the environmental pressure and performances of the energy sector. Gross Domestic Product (GDP) per capita, population density, and urbanization have been chosen as reliable indicators of affluence (A), coming from Eurostat. Data on GDP per capita are collected in chain-linked volume with 2010 as the basis year, whereas population density refers to the ratio of inhabitants per square kilometer. Urbanization measures the (urban) population percentage living in the largest city (e.g., metropolis) taken from World Bank’s World Development Indicators database. Furthermore, the algorithm will be fed with data on industrialization, i.e., percentage of manufacturing Gross Value Added (GVA) in total GDP. The dataset covers a decade (2008-2018) for 31 European Countries (27 EU + Norway, United Kingdom). For the energy dimension, the choice of the variables encompasses significant aspects of the energy transition. Primary energy consumption and share of renewable energies have been chosen as measures of the two aspects of energy transition (i.e., energy efficiency, clean technologies). The two variables have been retrieved from the SDG Indicators collected by Eurostat. The energy dimension is completed with data on the shares of electricity production from (liquid and solid) fossil fuels computed from Eurostat. Over a decade, the clustering algorithm will be performed on three reference years (2008, 2013, 2018). Considering the different units of measures, the clustering algorithm will be performed on the variables after standardization. The analysis will be eventually conducted on a comparative static of the three years, highlighting changes in the composition of clusters. As for the clustering algorithm *partitioning around medoids* (PAM) has been chosen for better dealing with outliers, as suggested by Kaufman and Rousseeuw (2005). In this procedure, clusters are constructed by assigning each dataset object to the nearest representative object (e.g., *medoids*). Choosing the optimal number of clusters for all the years will follow a two-layered approach. The optimal number of clusters will be assessed via the elbow method as the first step. The choice of this method, with respect to others, is mainly driven by the capacity of grouping clusters in a more meaningful way, namely reducing the sum of squared errors. However, as there exist a quite wide array of methods (Milligan and Cooper 1985; Cuevas et al. 2000; Tibshirani et al. 2001), the choice will also follow a somehow graphical approach considering the more homogeneous configuration of clusters (Kaufman and Rousseeuw 2005).

## 5.1 Results and Discussion

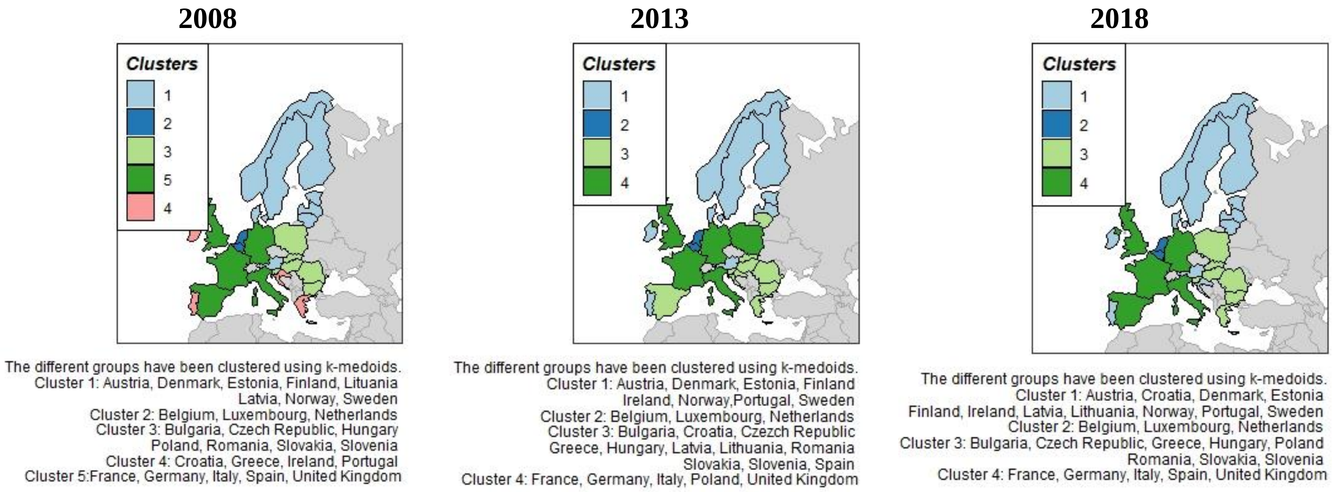
Figure 1 shows the elbow plot for the data on the three reference periods (2008, 2013, 2018). The optimal number of clusters *k* will be chosen whenever more disaggregated groups do not significantly reduce the sum of the squares within clusters. This stands for k = 5 in 2008, k = 4 in 2013 and k = 4 in 2018.

Figure 1 Elbow chart for the three reference periods



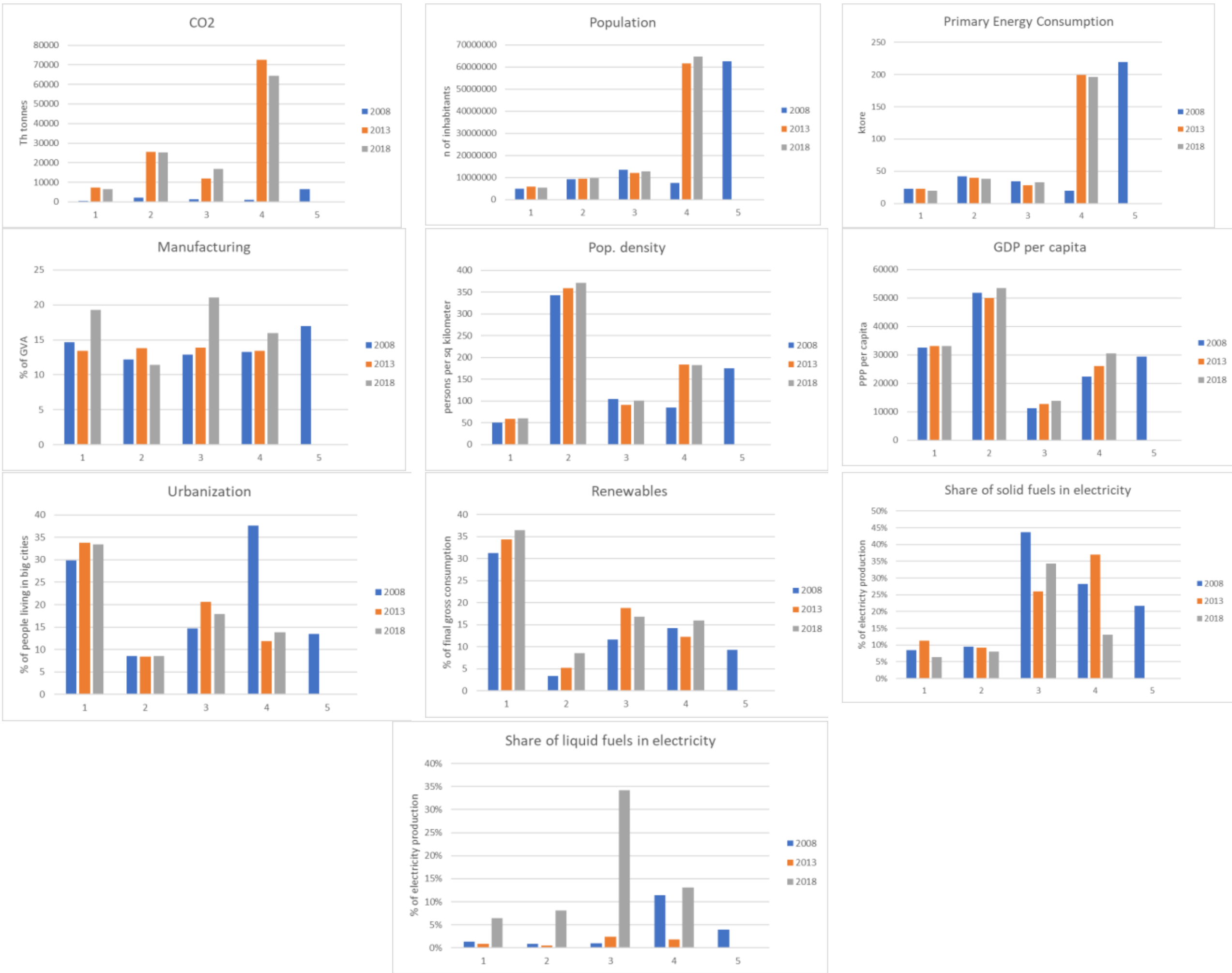

The analysis follows with the geographical representations of the clusters for the three reference periods (Figure 2). In this sense, as it was possible to appreciate from the preliminary analysis on the number of clusters, 2013 and 2018 show the same number of clusters. The clustering and the different average performances of countries contribute to creating a diversified picture over the years.

Figure 2 Groups of EU Countries clustered with K-medoids algorithm




From a preliminary overview, there appear to be groups of countries that have been persistent over time for the three different years. On the other hand, over the three years, there have been some changes in the composition of clusters. The analysis will focus on the formation and characteristics of every single year. Figure 3 summarizes the average quantities of the single variables for each cluster.

Figure 3 Average values for every cluster in the three reference years




*2008*

In Cluster 1 is grouped mostly northern European countries and Austria and Estonia. They are characterized for the highest average performances in terms of clean energy consumption (31 % with respect to 3% of Cluster 2, see Figure 3). This is coupled with the lowest emissions and primary energy consumption. Regarding the socioeconomic variables, Cluster 1 presents relatively high urbanization with low density with the second highest GDP per capita level. Cluster 2 groups Belgium, Luxembourg, and the Netherlands, the richest cluster per number of inhabitants with the highest density. As for environmental performances, Cluster 2 holds the second-highest share of emissions and the lowest share of renewable energies. Cluster 3 gathers all the Balkan, Baltic, and Eastern European States except Croatia and Estonia. This Cluster presents lower socioeconomic conditions and a higher density of inhabitants (105 people/sq meter on average see Figure 3). Those countries present the highest share of solid fuels employed in electricity production (43%, see Figure 3) and relatively low environmental pressure. On the other hand, the Cluster presents a relatively high performance in clean energy production (11% as opposed to 3% of Cluster 2) and energy efficiency. Croatia, along with Greece, Ireland, Portugal, form Cluster 4. This group displays the highest urbanization level (37%, see Figure 3) but the lowest density and number of inhabitants on average. From the energy side, it is the cluster with the second-highest share of renewables in the energy mix (14%). Still, it holds the highest shares of solid and liquid fossil fuels in electricity production (28% and 11%, respectively). Cluster 5 collects France, Germany, Italy, Spain, Poland with the highest share of carbon dioxide emissions on average. The cluster shows the highest energy consumption and the second-lowest share of renewables. It is the most populated cluster with relatively low urbanization but high density. Cluster 4 presents the higher share of manufacturing GVA in GDP.

*2013*

Cluster 1 groups the same countries as in 2008 with the addition of Ireland and Portugal (Figure 2). In fact, in 2013, the cluster maintained the same position as the best performer in energy consumed from renewable sources. The cluster shows the highest level of urbanization (33%, see Figure 3) the lowest density (59 inhabitants/sq km). Cluster 2 has not changed in composition from the previous. Inhabitants in those countries are the wealthiest, with the lowest share of clean energy. The cluster also presents the second-worst performance in terms of environmental pressure (e.g., $CO_2$ emissions). Cluster 3 (Bulgaria, Croatia, Czech Republic, Greece, Hungary, Romania, Latvia, Lithuania, Slovak, Slovenia, Spain) holds the lowest wealth per capita and the second-highest share of solid fuels in electricity production (25%, see Figure 3). On the other hand, the cluster presents relatively low carbon dioxide emissions coupled with a sustained share of renewable energies (18%). Cluster 4 in 2013 includes France, Germany, Italy, Poland, United Kingdom. It shows the highest level of emissions with the highest level of primary energy consumption. Furthermore, Cluster 4 holds the highest share of electricity production from (solid) fossil fuels (37% see Figure 3). It also appears to be the highest populated cluster despite the relatively low level of urbanization.

*2018*

Cluster 1 keeps all the countries as in 2013 with Croatia, Latvia, and Lithuania. It is the least populated, also in terms of density. This group of Countries is the second wealthy cluster in terms of GDP per capita. Moreover, Cluster 1 is the lowest emitter of carbon dioxide in the atmosphere and the highest energy consumer from renewable sources. Despite the relatively high contribution of the industrial sector, the cluster appears to consume a quite low quantity of energy (e.g., Primary Energy Consumption). Cluster 2 (Belgium, the Netherlands, Luxembourg) represents the wealthiest and the most densely populated agglomeration of EU States. On the other hand, it holds the lowest share of renewables (8% Figure 3) and a low industrial productivity level. Cluster 3 gathers Bulgaria, Czech Republic, Greece, Hungary, Romania, Slovak, Slovenia. It holds the lowest GDP per inhabitant but the highest share of industry's contribution to the GDP (21% see Figure 2). On average, 17% of people live in the biggest cities, and the primary energy consumption is the lowest with respect to the other clusters. Cluster 3 holds the highest share of electricity production from (solid and liquid) fossil fuels, reaching 34% of electricity produced in 2018. On the other hand, this cluster shows 33% of energy coming from renewable sources (the highest percentage, 38%, pertains to Cluster 1). Cluster 4 groups France, Germany, Italy, Spain, and the United Kingdom with the highest population level on average. The cluster appears to be the highest emitter of $CO_2$ in the atmosphere and the highest energy consumer. Despite the highest population level, this cluster shows relatively low levels of both density and urbanization.

## 5.2 Discussion

The analysis of the clusters has identified groups of countries with peculiar characteristics. In this sense, Cluster 1 has always gathered countries that have shown relatively good performance in clean energy consumption, energy efficiency, and carbon dioxide emissions. Countries in that cluster have always been characterised by not being higher populated or densely inhabited. The solid basis of the cluster has been represented by Nordic countries with similar socioeconomic, cultural, and demographic characteristics Blindheim (2015). Austria shares the same performances, whereas it holds a higher population density due to the physical territorial extension of the States. Latvia and Lithuania were present in 2008 and 2018. On the other hand, Croatia, Ireland, and Portugal joined Cluster 1 in 2018. Latvia has shown an increasing share of renewables and relatively low energy consumption patterns. The Investment and Development Agency of Latvia (2020) reported that the Country is now in third place among the EU countries regarding renewable energy consumption. Together with Estonia and Lithuania, those two Baltic countries have shown good renewable production and consumption performances. The favourable climatic condition also helped Latvia, whereas Estonia still appears to be highly dependent on carbon-based energy (Štreimikiene et al., 2016). Croatia managed to increase its overall energy efficiency by 21.4% in the period 2000-2018, mostly led by industry (+ 2.4% per year) and the residential sector (+1.4% per year) (Odyssee ,2021). As for Portugal, Østergaard et al. (2014) observe the Country has used a progressively consistent amount of renewables in the energy mix with the help of favourable climatic conditions to combat energy dependence. Ireland represents the latest newcomer to the Cluster in this picture in terms of (clean) energy performances. Although the share of renewables in consumption has increased over time (4% in 2008, 13% in 2013, 10.88% in 2018), it is still the lowest share. However, energy consumption patterns appear to be in line with the other countries of the cluster. Cluster 2 has been stable for all the periods considered with Belgium, Luxembourg, and the Netherlands. As already disentangled, those countries share the same socioeconomic characteristics with high per capita wealth and a relatively small territorial extension. Indeed, countries in the cluster appear to be the most densely populated. As for energy performances, Cluster 2 shows the smallest share of renewables yet with an increasing trend (up to 8.6% on average in 2018). Quantities of energy consumption have slightly decreased over time along with industrial performances. In 2008 there was also the formation of Cluster 4 with Croatia, Greece, Ireland, Portugal. Greece joined Cluster 3 in 2013 and has not moved since then. On the other hand, Croatia, Ireland, Portugal have joined Cluster 1 since 2013. Newcomers in Cluster 1 in 2018 are mostly related to increasing clean energy consumption and performance efficiency. As already argued, Cluster 3 shows lower economic wealth per capita with an increasing level of urbanization. The cluster also presents a low level of primary energy consumption coupled with lower carbon dioxide levels from fuel combustion. According to the estimates by World Energy Council and Oliver Wyman (2020), Romania scores among the countries with the highest capacity of meeting energy demand internally (e.g., energy security). However, Romania still benefits from being an oil producer while it is still in the process of applying the EU energy agenda. In this sense, in the past two decades, the share of fossil fuel in the energy supply in the country has decreased and replaced with renewables (+10% within 2000-2018) and nuclear power (+5-6% within 2000-2018) (World Energy Council and Oliver Wyman, 2020). In Hungary, the share of nuclear energy accounted for 37% of total final consumption (TFC) in 2015, thus covering the decrease in fossil fuels[2] (IEA 2017b). As for Greece, the country has heavily relied on coal (i.e., lignite) production and imports of oil with a small but increasing share of renewables (mostly biofuel and waste) (IEA, 2017a). Reports by Agency of Energy (2019); Ministry of Environment (2018) show Slovenia and Slovakia can be classified as net energy importers with a high share of nuclear power in internal generation. Poland has changed position from Cluster 3 in 2008 and 2018 to Cluster 4 in 2013. Poland has heavily relied on fossil fuel, especially coal, for its energy mix (74.4% of electricity generation from coal in 2020) (Hasterok et al., 2021). Despite the pressing influence of the EU environmental objective of a net-zero economy by 2050, the Polish government still plans to rely on fossil fuels for a long time (Kudełko, 2021). However, according to Polish Ministry of Climate and Environment (2021), Poland will be introducing nuclear power plants in its energy mix by the third decade of 2000 to lower the incidence of coal sources. The position in the cluster with Germany, France, Italy, United Kingdom has probably been achieved due to their bad performances in terms of emissions and clean energy consumption relative to the size of the economy.

France, Germany, Italy, United Kingdom have formed another stable bloc of countries over time. In 2008 and 2018, those four countries were joined by Spain. In 2013 Spain was replaced by Poland in the cluster. This replacement is mainly related to the 2008 economic crisis that particularly hit the Spanish economy in 2013[3]. Countries of this group have scored marked (worse) environmental performances with higher $CO_2$ emissions coupled with a high population level. From the energy side, Countries in clusters 4-5 show higher energy consumption and relatively low but increasing performances in renewable consumption (from an average of 9% in 2008 to 15% in 2018). Despite this higher energy consumption with respect to the other Countries, this cluster shows average industrial sector performance levels over time. However, Alola et al. (2019) proved that carbon dioxide and housing positively impact renewable energy generation in the long run, especially for Mediterranean countries. The cluster contains the most developed economies of the Union, and despite their commitment to EU objectives, they still appear to lag in clean energy generation. According to Telli et al. (2021), the reason can be tracked down to the lack of available space to implement renewable energy generation for the national demand for energy or reliance on other sources for energy production (e.g., France, Spain). Indeed, despite the high commitment of Germany, the country still heavily relies on fossil fuels for energy supply (80% of primary energy supply in 2018). In contrast, for France, nuclear energy contributed 46.6% in 2015 (IEA 2016a, 2020). According to the latest data by IEA (2021a), nuclear energy covers around

[2] In 2015 the country has gone from self-sufficient to being dependent for 87% on imports of natural gas. The same share stands for crude oil. However, the country still relies on coal for two-thirds of TFC

[3] https://www.expansion.com/2013/12/18/economia/1387360918.html

45% of production in Spain. The case of Italy[4] is different with an increase in production due to renewable energies (68% in 2015) despite total energy supply still heavily relying on fossil fuels (IEA 2016b).

## 6. Conclusions

Since the beginning of the European Union, the so-called harmonization process has paved the way to a common orientation for the Member States, leaving broad discretion to each national regulatory framework (Majone 2014). The EU regulatory framework is one of the most stringent and comprehensive globally. The Union is among the leaders and signatories of many international agreements (Paris Agreement, COP on Climate Change, COP on Biodiversity). In this framework, the latest roadmap at the policy level, the EU has committed to reaching net-zero carbon emissions by 2050 agenda series of other objectives for clean energy, energy efficiency, biodiversity, ecosystem conservation, sustainable production, and consumption (EU Green Deal). This overall (financial and non) impulse will ideally improve Member States' commitments and environmental performances. Ideally, the European Union should act as a cohesive entity in the international landscape. Cultural and historical differences also mark consistent divergences within the Member States regarding environmental policies and the overall orientation of the EU (Jehlička and Tickle 2004). Applying a data-driven approach, this work tries to provide a comprehensive picture of how the Member States are coping with their environmental commitments applying a consolidated analytical framework. The IPAT relationship provided an overarching analytical setting to assess environmental performances comprising social, technological, and economic aspects. However, cluster analysis is highly dependent on the choice of the data to feed the algorithm. This work has focused on those variables that have been traditionally employed in the analysis of the IPAT identity in the literature. The choice of this approach was to provide a framework to the convergence of EU policy that comprises all the relevant dimensions (i.e., economic, demographic, environmental). Further expansions of the analysis may envision a set of variables more in line with the EU's objectives in terms of the low-carbon transition. A clustering algorithm has been applied to three cross-sections of data on three different periods (2008,2013,2018). The analysis identified three specific groups with marked differences: the ones with higher performances in terms of clean energy, energy efficiency; wealthy countries with poor environmental performances instead of relatively poorer countries with promising environmental performances. Among those polarized clusters, some countries have moved through clusters. After a transition phase (2013) in 2018, Latvia, Lithuania joined the cluster of best environmental performers. Croatia, Ireland, Portugal managed to reach Cluster 1 in 2018. On the other hand, Poland joined the cluster of bad environmental performers in 2013. On the other hand, Spain joined France, Germany, Italy, and the United Kingdom in 2008 and 2018. What marks a consistent divergence with the other clusters is the demographic (P) dimension. Even though clustering does not allow for causal relationships, it is possible to affirm that population size plays a consistent role in a country's environmental performance. The bulk of the policy framework at the EU level on air emissions and energy efficiency addresses heavy polluting industrial sectors (e.g., energy, petrochemical, industry). On the other hand, residential heating and cooling represent 46% of the total energy consumption for heating and cooling[5] (IEA 2021b). The EU Directive 2012/27/EU (e.g., Energy Efficiency Directive) implements specific measures to promote a precise account of energy consumption related to heating and cooling within (non-)residential buildings meeting the overall energy efficiency targets at the EU level. In this sense, despite all the incentives to promote a more renewable-oriented mix for heating and cooling (e.g., building energy codes), this energy consumption side is overlooked in the environmental policy framework according to IEA (2021b). Barriers to implementing that kind of clean technologies in (non-) residential buildings can be tracked down to the difficulty of the payback mechanism the difficulty of reaching an agreement on the investment (i.e., residential buildings). On the other hand, Fraunhofer Institute for Systems and Innovation Research et al. (2017) find investments for non-residential buildings often are undertaken if they provide concrete advantages in labour productivity (e.g., a better work environment for employees). In fact, of all European Member States, only Croatia has not set specific policy options for heating and cooling in any sector (REN21, 2021). Another interesting emerging pattern is nuclear energy within the energy mix as a substitute for fossil fuel-based sources. Nuclear power represents more than 50% of electricity consumption

[4] Following the results of the referendum in 1987 the Country decommissioned all the nuclear power stations abolishing nuclear energy from its energy mix.

[5] 72% of energy consumption for heating and cooling comes from coal sources IEA (2021b)

in France, Slovak, and Hungary and is a higher low-carbon source for other EU Countries (IEA, 2019a). Nuclear power is still under consideration to be included in a Delegated Act of the EU Taxonomy. Indeed, it has been argued that the technology meets the "do-no-significant-harm" (DNSH) principle. European Commission Joint Research Center (2021); Scientific Committee on Health, Environmental and Emerging Risks (2021); IAEA (2021) highlight the main critical points revolve around (hazardous) waste production and material efficiency of existing plants in the use of uranium[6], considering its extraction and the rather insufficient attention to the impact of radiation on (marine) ecosystems.

Overall, thiw work tries to shed light on the state of convergence of national patterns in environmental policy implementation, meeting the objectives at the EU level. Moreover, the work tries to provide a new perspective of employment of a well-established analytical framework. Clustering techniques allow systematization and classification via a sole data-driven approach without any inference on the relationship among variables. Furthermore, the results of cluster analyses are highly dependent on the number and nature of the variables considered. Despite those limitations, the work feeds the literature of policy convergence, providing a systematic classification of countries with respect to their performances in relevant areas of policy intervention. The comparative static analysis over three reference periods provided a diverse picture of environmental performance. The choice of the data has primarily followed the literature on IPAT analysis to provide an analytical framework encompassing all the dimensions of the human-environmental relationship (i.e., economic, demographic, technological, environmental). Further development of this work might include variables more directly related to the objectives set at the EU level on low-carbon transition. Overall, the landscape of the EU Member States presents persistent clusters of Countries over time, also considering the socioeconomic dimension. In fact, after a transition phase (for most of the States in 2013), some Countries managed to increase their performances in terms of clean energy consumption, emission reduction, energy efficiency, whereas some others did not. Ideally, the ultimate aim of the European Union would be to harmonize the Member States in terms of policymaking and (environmental) outcomes; the analysis suggests much work has already been done, whereas much more is needed.

## References


Agency of Energy. 2019. 'REPORT ON THE ENERGY SECTOR IN SLOVENIA'. Agency of Energy-Slovenia. https://www.agen-rs.si/documents/54870/68629/Report-on-the-energy-sector-in-Slovenia-for-2019/ce1c3cd8-489a-401d-9a1a-502a7c5715e4.

Alola, Andrew Adewale, Uju Violet Alola, and Seyi Saint Akadiri. 2019. 'Renewable Energy Consumption in Coastline Mediterranean Countries: Impact of Environmental Degradation and Housing Policy'. *Environmental Science and Pollution Research* 26 (25): 25789–801. https://doi.org/10.1007/s11356-019-05502-6.

Babenko, V., A. Panchyshyn, L. Zomchak, M. Nehrey, Z. Artym-Drohomyretska, and T. Lahotskyi. 2021. 'Classical Machine Learning Methods in Economics Research: Macro and Micro Level Examples'. *WSEAS Transactions on Business and Economics* 18: 209–17. https://doi.org/10.37394/23207.2021.18.22.

Banga, Alisha, and Amrita Sinha. 2018. 'Clustering Application for Streaming Big Data in Smart Grid'. In *2018 International Conference on Communication and Signal Processing (ICCSP)*, 1051–54. https://doi.org/10.1109/ICCSP.2018.8524505.

Blindheim, Bernt. 2015. 'A Missing Link? The Case of Norway and Sweden: Does Increased Renewable Energy Production Impact Domestic Greenhouse Gas Emissions?' *Energy Policy* 77 (February): 207–15. https://doi.org/10.1016/j.enpol.2014.10.019.


[6] The current technology of nuclear plants implies a high percentage of activated uranium not recyclable. Recent technological development (e.g., *fast-neutron spectrum*) will imply the exploitation of the material 50 times higher than the current rate (European Commission Joint Research Center, 2021)

Boumans, Marcel, and Sabina Leonelli. 2020. ‘From Dirty Data to Tidy Facts: Clustering Practices in Plant Phenomics and Business Cycle Analysis’. In *Data Journeys in the Sciences*, edited by Sabina Leonelli and Niccolò Tempini, 79–101. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-37177-7_5.

Brizga, Janis, Kuishuang Feng, and Klaus Hubacek. 2013. ‘Drivers of CO2 Emissions in the Former Soviet Union: A Country Level IPAT Analysis from 1990 to 2010’. *Energy* 59 (September): 743–53. https://doi.org/10.1016/j.energy.2013.07.045.

Cervone, G., P. Franzese, Y. Ezber, and Z. Boybeyi. 2008. ‘Risk Assessment of Atmospheric Emissions Using Machine Learning’. *Natural Hazards and Earth System Sciences* 8 (5): 991–1000. https://doi.org/10.5194/nhess-8-991-2008.

Chang, Yi Ling, and Yuan-Yuan Lee. 2019. ‘Applied Machine Learning: A Creation of Sustainable Development Progress Index Based on Circular Economy’. In *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, 1263–68. https://doi.org/10.1109/CSCI49370.2019.00237.

Chen, Shang-Yuan, and Tzu-tien Chen. 2020. ‘Application of Cluster Analysis with Unsupervised Learning to Dockless Shared Bicycle Flow Control and Dispatching’. *Computer-Aided Design and Applications* 17 (5): 1067–83. https://doi.org/10.14733/cadaps.2020.1067-1083.

Chertow, Marian R. 2000. ‘The IPAT Equation and Its Variants’. *Journal of Industrial Ecology* 4 (4): 13–29. https://doi.org/10.1162/10881980052541927.

Chontanawat, Jaruwan. 2018. ‘Decomposition Analysis of CO2 Emission in ASEAN: An Extended IPAT Model’. *Energy Procedia*, 5th International Conference on Energy and Environment Research, ICEER 2018, 23-27 July 2018, Prague, Czech Republic, 153 (October): 186–90. https://doi.org/10.1016/j.egypro.2018.10.057.

Cuevas, Antonio, Manuel Febrero, and Ricardo Fraiman. 2000. ‘Estimating the Number of Clusters’. *Canadian Journal of Statistics* 28 (2): 367–82. https://doi.org/10.2307/3315985.

DG Energy. 2020. ‘EU Energy in Figures’. European Commission. https://op.europa.eu/en/publication-detail/-/publication/87b16988-f740-11ea-991b-01aa75ed71a1/language-en?WT.mc_id=Searchresult&WT.ria_c=37085&WT.ria_f=3608&WT.ria_ev=search.

Di, Zhenzhen, Miao Chang, Peikun Guo, Yang Li, and Yin Chang. 2019. ‘Using Real-Time Data and Unsupervised Machine Learning Techniques to Study Large-Scale Spatio–Temporal Characteristics of Wastewater Discharges and Their Influence on Surface Water Quality in the Yangtze River Basin’. *Water* 11 (6): 1268. https://doi.org/10.3390/w11061268.

Dogan, A., and D. Birant. 2021. ‘Machine Learning and Data Mining in Manufacturing’. *Expert Systems with Applications* 166. https://doi.org/10.1016/j.eswa.2020.114060.

Eggleston, H. S, Intergovernmental Panel on Climate Change, National Greenhouse Gas Inventories Programme, and Chikyū Kankyō Senryaku Kenkyū Kikan. 2006. *2006 IPCC Guidelines for National Greenhouse Gas Inventories*. Vol. 2. http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.htm.

Ehrlich, Paul R., and John P. Holdren. 1971. ‘Impact of Population Growth’. *Science* 171 (3977): 1212–17.

European Commission. 2020. ‘Country Report Lithuania 2020’. COMM(150). 2020 European Semester: Assessment of Progress on Structural Reforms, Prevention and Correction of Macroeconomic Imbalances, and Results of in-Depth Reviews under Regulation (EU) No 1176/2011. Bruxelles: European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020SC0514&from=EN.

European Commission Joint Research Center. 2021. ‘Technical Assessment of Nuclear Energy with Respect to the “Do No Significant Harm” Criteria of Regulation (EU) 2020/852 (“Taxonomy Regulation”)’. Joint

Research Center. https://ec.europa.eu/info/sites/default/files/business_economy_euro/banking_and_finance/documents/210329-jrc-report-nuclear-energy-assessment_en.pdf.

Franceschi, Fabiana, Martha Cobo, and Manuel Figueredo. 2018. ‘Discovering Relationships and Forecasting PM10 and PM2.5 Concentrations in Bogotá, Colombia, Using Artificial Neural Networks, Principal Component Analysis, and k-Means Clustering’. *Atmospheric Pollution Research* 9 (5): 912–22. https://doi.org/10.1016/j.apr.2018.02.006.

Fraunhofer Institute for Systems and Innovation Research, Fraunhofer Institute for Solar Energy Systems, Institute for Resource Efficiency and Energy Strategies GmbH, Observ’ER, Technical University Vienna - Energy Economics Group, and TEP Energy GmbH. 2017. ‘Mapping and Analyses of the Current and Future (2020 - 2030) Heat-Ing/Cooling Fuel Deployment (Fossil/Renewables)’. Final report. European Commission. https://ec.europa.eu/energy/sites/default/files/documents/mapping-hc-final_report-wp5.pdf.

Hasterok, Damian, Rui Castro, Marcin Landrat, Krzysztof Pikoń, Markus Doepfert, and Hugo Morais. 2021. ‘Polish Energy Transition 2040: Energy Mix Optimization Using Grey Wolf Optimizer’. *Energies* 14 (2): 501. https://doi.org/10.3390/en14020501.

IAEA. 2021. *URANIUM RAW MATERIAL FOR THE NUCLEAR FUEL CYCLE: Exploration, Mining, Production, Supply and... Demand, Economics and Environmental Issues.* S.l.: INTL ATOMIC ENERGY AGENCY.

IEA. 2016a. ‘Energy Policies of IEA Countries - Italy 2016 Review’. *Energy Policies of IEA Countries*, 214.

———. 2016b. ‘Energy Policies of IEA Countries France 2016 Review’. *Energy Policies of IEA Countries*, 211.

———. 2017a. ‘Energy Policies of IEA Countries - Greece Review 2017’, 143.

———. 2017b. ‘Energy Policies of IEA Countries - Hungary 2017 Review’, 176.

———. 2019. ‘Nuclear Power in a Clean Energy System’, 103.

———. 2020a. ‘CO2 Emissions from Fuel Combustion: Overview 2020’, 13.

———. 2020b. ‘Germany 2020 - Energy Policy Review’, 229.

———. 2021a. ‘Energy Policy Review- Spain 2021’. International Energy Agency. https://iea.blob.core.windows.net/assets/2f405ae0-4617-4e16-884c-7956d1945f64/Spain2021.pdf.

———. 2021b. ‘Renewable Energy Policies in a Time of Transition: Heating and Cooling’, 150.

Investment and Development Agency of Latvia. 2020. ‘Environment and Renewable Energy Industry.Pdf’. https://www.liaa.gov.lv/en/trade/industries/environment-and-renewable-energy.

IPCC. 2015. *IPCC Special Report: Global Warming of 1.5°*. IPCC Special Report. https://www.ipcc.ch/site/assets/uploads/sites/2/2019/05/SR15_Chapter1_Low_Res.pdf.

Jayatilake, S.M.D.A.C., and G.U. Ganegoda. 2021. ‘Involvement of Machine Learning Tools in Healthcare Decision Making’. *Journal of Healthcare Engineering* 2021. https://doi.org/10.1155/2021/6679512.

Jehlička, Petr, and Andrew Tickle. 2004. ‘Environmental Implications of Eastern Enlargement: The End of Progressive EU Environmental Policy?’ *Environmental Politics* 13 (1): 77–95. https://doi.org/10.1080/09644010410001685146.

Kaufman, Leonard, and Peter J. Rousseeuw. 2005. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Mathematical Statistics. Hoboken, N.J: Wiley.

Kudełko, Mariusz. 2021. ‘Modeling of Polish Energy Sector – Tool Specification and Results’. *Energy* 215 (January): 119149. https://doi.org/10.1016/j.energy.2020.119149.

Majone, Giandomenico. 2014. ‘Policy Harmonization: Limits and Alternatives’. *Journal of Comparative Policy Analysis: Research and Practice* 16 (1): 4–21. https://doi.org/10.1080/13876988.2013.873191.

McNicoll, Geoffrey. 2015. ‘IPAT (Impact, Population, Affluence, and Technology)’. In *International Encyclopedia of the Social & Behavioral Sciences*, 716–18. Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.91045-6.

Milligan, Glenn W, and Martha C Cooper. 1985. ‘An Examination of Procedures for Determining the Number of Clusters in a Data Set’. *Phychometrica* 50 (2): 159–79.

Ministry of Environment. 2018. ‘STATE OF THE ENVIRONMENT REPORT – SLOVAK REPUBLIC’. Ministry of Environment of the Slovak Republic. https://www.enviroportal.sk/uploads/report/10538.pdf.

Østergaard, Poul Alberg, Isabel Soares, and Paula Ferreira. 2014. ‘Energy Efficiency and Renewable Energy Systems in Portugal and Brazil’. *International Journal of Sustainable Energy Planning and Management* 2 (June): 1–6. https://doi.org/10.5278/ijsepm.2014.2.1.

Polish Ministry of Climate and Environment. 2021. ‘OBWIESZCZENIE MINISTRA KLIMATU I ŚRODOWISKA z dnia 2 marca 2021 r. w sprawie polityki energetycznej państwa do 2040 r.’, 568.

Quatrosi, Marco. 2020. ‘Analysis of Monthly CO2 Emission Trends for Major EU Countries: A Time Series Approach’. *SEEDS Working Paper Series*, no. 15: 25.

REN21. 2021. ‘Renewables Global Status Report 2021’. REN 21. https://www.ren21.net/wp-content/uploads/2019/05/GSR2021_Full_Report.pdf.

Sands, Philippe, and Paolo Galizzi, eds. 2006. ‘Decision No 1600/2002/EC of the European Parliament and of the Council of 22 July 2002 Laying down the Sixth Community Environment Action Programme (OJ L 242 10.09.2002 p. 1)’. In *Documents in European Community Environmental Law*, 2nd ed., 63–89. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610851.007.

Scientific Committee on Health, Environmental and Emerging Risks. 2021. ‘SCHEER Review of the JRC Report on Technical Assessment of Nuclear Energy with Respect to the “Do No Significant Harm” Criteria of Regulation (EU) 2020/852 (“Taxonomy Regulation”)’, 16.

Silva, Benedito Albuquerque da, Michel Constantino, Ozeni Souza de Oliveira, Sandro Aparecido Lima dos Santos, Benjamin Miranda Tabak, and Reginaldo Brito da Costa. 2019. ‘New Indicator for Measuring the Environmental Sustainability of Publicly Traded Companies: An Innovation for the IPAT Approach’. *Journal of Cleaner Production* 215 (April): 354–63. https://doi.org/10.1016/j.jclepro.2019.01.039.

Steinley, Douglas. 2006. ‘K-Means Clustering: A Half-Century Synthesis’. *British Journal of Mathematical and Statistical Psychology* 59 (1): 1–34. https://doi.org/10.1348/000711005X48266.

Štreimikiene, D., W. Strielkowski, Y. Bilan, and I. Mikalauskas. 2016. ‘Energy Dependency and Sustainable Regional Development in the Baltic States - A Review’. *Geographica Pannonica* 20 (2): 79–87. https://doi.org/10.5937/GeoPan1602079S.

Telli, Azime, Selma Erat, and Bunyamin Demir. 2021. ‘Comparison of Energy Transition of Turkey and Germany: Energy Policy, Strengths/Weaknesses and Targets’. *Clean Technologies and Environmental Policy* 23 (2): 413–27. https://doi.org/10.1007/s10098-020-01950-8.

Tibshirani, Robert, Guenther Walther, and Trevor Hastie. 2001. ‘Estimarting the Number of Clusters in a Data Set via the Gap Statistic’. *Journal of Royal Statistical Society* B (2): 411–23.

Wen, Lei, and Zhenkai Li. 2019. ‘Driving Forces of National and Regional CO2 Emissions in China Combined IPAT-E and PLS-SEM Model’. *Science of The Total Environment* 690 (November): 237–47. https://doi.org/10.1016/j.scitotenv.2019.06.370.

World Energy Council, and Oliver Wyman. 2020. ‘World Energy Trilemma Index 2020’. World Energy Council. https://trilemma.worldenergy.org/reports/main/2020/World%20Energy%20Trilemma%20Index%202020.pdf.

Yue, Ting, Ruyin Long, Hong Chen, and Xin Zhao. 2013. ‘The Optimal CO2 Emissions Reduction Path in Jiangsu Province: An Expanded IPAT Approach’. *Applied Energy* 112 (December): 1510–17. https://doi.org/10.1016/j.apenergy.2013.02.046.