

HOSTED BY



Contents lists available at ScienceDirect

Atmospheric Pollution Research

journal homepage: <http://www.journals.elsevier.com/locate/apr>

Original article

Research of varying levels of greenhouse gas emissions in European countries using the k-means method



Anna Kijewska, Anna Bluszczyk*

Department of Mining Management and Safety Engineering, Silesian University of Technology, Poland

ARTICLE INFO

Article history:

Received 20 December 2015

Accepted 20 May 2016

Available online 9 June 2016

Keywords:

Greenhouse gas emissions

GHG

Cluster analysis

k-means method

ABSTRACT

Greenhouse gas emissions are a global problem. Although the EU countries from 1990 to 2012 reduced their total emissions by 19.2% (CO₂ eq.), it is still necessary to limit their emissions. In the article the possibility of using the taxonomic methods that allow grouping (classifying) objects described by many attributes (variables) is presented. In particular, cluster analysis was used, in which some methods for the isolation of homogeneous subsets of surveyed objects can be distinguished. One of such method is *k-means* algorithm. As a measure of similarity of objects in clusters the Euclidean distance was applied. In the analysis 28 European countries were taken as objects of research and they were described by four attributes (variables), i.e. the emission levels of greenhouse gases such as carbon dioxide, methane, nitrogen oxides and nitrous oxide. The aim of the analysis is to grouping objects – the European countries – into clusters that are most similar to each other in the same cluster and most unlike in other clusters. The research was carried out according to total greenhouse gas emissions, and according to emissions of these gases per capita of the countries surveyed. The analyses are based on Eurostat reports.

Copyright © 2016 Turkish National Committee for Air Pollution Research and Control. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Coal mining is a strategic sector of the Polish economy and plays a key role in ensuring the energy security of the country (Burchart-Korol et al., 2014). Poland ranks second in Europe in terms of volume of coal mining output and the fourth largest in terms of lignite mining output (GUS, 2014). Coal is, and will be in the future, the primary source of energy, as the share of coal fuel in electricity generation in Poland is nearly 90% (Dubiąski and Turek, 2014; PEP, 2014). State policy based on the principles of sustainable development is closely linked with the mining industry, which is the main pillar for the further development of society, growth of the life quality of the population and the economic development of Poland, as well as EU countries.

The extractive industry is one of the most significant sector in shaping the global economic situation in the world (Ranosz, 2014). This is confirmed by trends in the global market for mining of

minerals, because countries that are rich in natural resources such as Brazil, Russia, India and China were the fastest growing economies of the world that already in 2000–2008 accounted for 50% of global economic growth (Pitfield et al., 2010).

Since 1990 a steady increase in world production of mineral resources has been seen, which indicates the relationship between mining and economic development (Dubiąski, 2013). Despite limiting the role of coal in meeting energy needs in European Union countries, coal will still serve an important role in the world, and its share in energy production in 2035 will be at the level of 26.3% (EO, 2015). The share of coal in the global energy production by regions of the world is presented in Fig. 1.

Undisputed is the fact of the negative impact of the use of mineral resources on the environment. At the same time, however, it should be stressed the significant involvement of mining companies in the eco-friendly and pro-social activities (Pietrzyk-Sokulska et al., 2015; Majer, 2013). The activity in this field confirms publications of integrated reports containing financial-economic, environmental and social information, i.e. Corporate Social Responsibility (CSR) reports (Hąbek, 2012, 2014; Hąbek and Wolniak, 2016).

Burning fossil fuels causes emissions of pollutants into the atmosphere in the form of gas and dust. Poland by signing the United

* Corresponding author. Silesian University of Technology, 2 Akademicka St., 44-100, Gliwice, Poland. Tel./fax: +48 32 2371296.

E-mail address: anna.bluszczyk@polsl.pl (A. Bluszczyk).

Peer review under responsibility of Turkish National Committee for Air Pollution Research and Control.

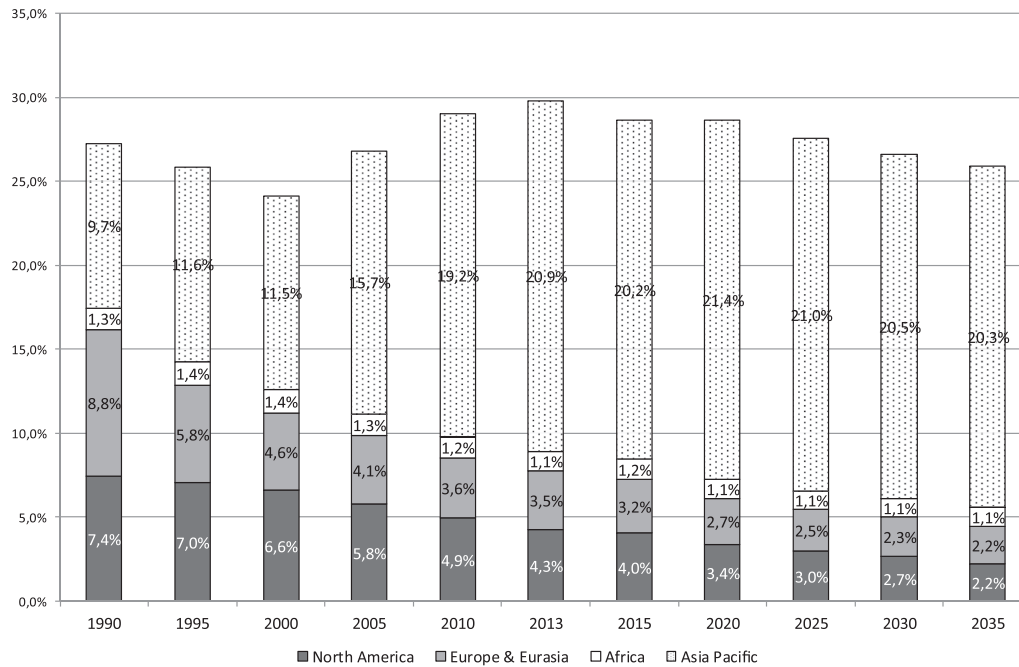


Fig. 1. Coal's share of world energy production from 1990 to 2035. Source: own work based on (EO, 2015).

Nations Framework Convention in 1994 and the Kyoto Protocol in 2002 committed to activities that would stabilize greenhouse gas (GHG) emissions (MŚ, 2003). Poland reduced its GHG emissions in 2012 as required by 14.15% compared to the base year – 1990 (IEA, 2014). Meeting the requirements for reducing CO₂ emissions became the basis for the development of multi-technology research in Carbon Capture and Storage (CCS) (Krziemień et al., 2013; Lutyński, 2014; Olajire, 2013; Uliasz-Bocheńczyk, 2010; Uliasz-Bocheńczyk and Mokrzycki, 2014; Uliasz-Bocheńczyk et al., 2009).

Gas emission limits in the EU can, however, be inadequate to protect the climate on a world scale, because the share of 28 EU countries in global emissions amounted to only 10.5% in 2013. Therefore, the Commission takes the view that to fight against global warming should join the fastest growing economy in the world such as China and India, which became the world's biggest emitters of greenhouse gases. Share of selected countries in global emissions in 1990 and 2013 is presented in Fig. 2 (NEAA, 2014).

In 2005, the developed countries accounted for about 40% of global CO₂ emissions, developing countries for about 56%, and the remaining 4% came from aviation and maritime transport, which emissions according to the internationally accepted methodology are not allocated to any particular region. At baseline, the developed countries will account in 2030 for about 32% of CO₂ emissions, developing countries for about 63%, and aviation and maritime transport for 5%. In 2030 it is expected that emissions per capita in developed countries will be still at more than double compared to the developing countries (respectively 16 and 7 t CO₂ eq. per year), although the expected increase of 0.7% per year in developed countries will be about $\frac{2}{3}$ lower than in developing countries (2.2% per year) (McKinsey, 2010).

At present in the scientific community still there is no clear opinion on the size of the human impact on global warming that is observed in recent years, and caused by the increase in greenhouse gas emissions. On the one hand the position is maintained that most of the observed increase in global average temperatures since the mid-twentieth century is with a high probability due to the observed increase in anthropogenic greenhouse gas concentrations (Pachauri and Meyer, 2014), which is reflected in activities on emission limits

for EU countries. On the other hand, there are differing opinions, undermine this approach (Idso et al., 2013) and emphasizing the complexity of this issue, among others, for the sake of the fact that burning fossil fuels is only one of many factors affecting the level of emissions of GHG such as e.g. a change in land use with comparable effect. Therefore, there is an opinion that unilateral approach to combating global warming by introducing restrictive CO₂ emission limits is not fully justified (Pawłowski and Cao, 2014).

Scientists in their research are looking for the sources and factors affecting the GHG emissions on the one hand, and on the other they are searching the methods and strategies that could lead to a reduction of these emissions. Among numerous publications on this topic it is worth mentioning (Xia and Chen, 2012), where five scenarios of energy abatement were evaluated, which were based on the combinations of regulation strategies and allocation alternatives. The authors, among other things come to the conclusion that the sector regulation strategy is more effective than the regional regulation strategy. In turn, in Xia et al. (2015) structural decomposition analysis is used to find the origins of changes of environmental and economic variables and key factors that have the greatest influence on these changes. Fuzzy cluster analysis is conducted in Xia et al. (2011) to group industrial sectors (in China), taking into account several indicators from the point of view of energy security, efficiency and carbon emission. Li et al. (2016) suggest that such strategies as displacing energy-intensive sectors to other regions lead to "local reduction but overall rise". They analyse five energy abatement scenarios in order to find the optimal abatement scenario for Beijing. Of course there is no certainty that such a scenario will work, for example in Europe, but carried out "three-scale input-output" analysis can certainly be used for other regions.

The aim of the article is to analyse the differences in the levels of GHG such as carbon dioxide, nitrogen oxides, methane, and nitrous oxides in European countries. These four groups of gases are considered to be the greatest threat, and only data on emissions of these gases are published by Eurostat. The result of the research is to identify clusters of similar countries, i.e. homogeneous objects in

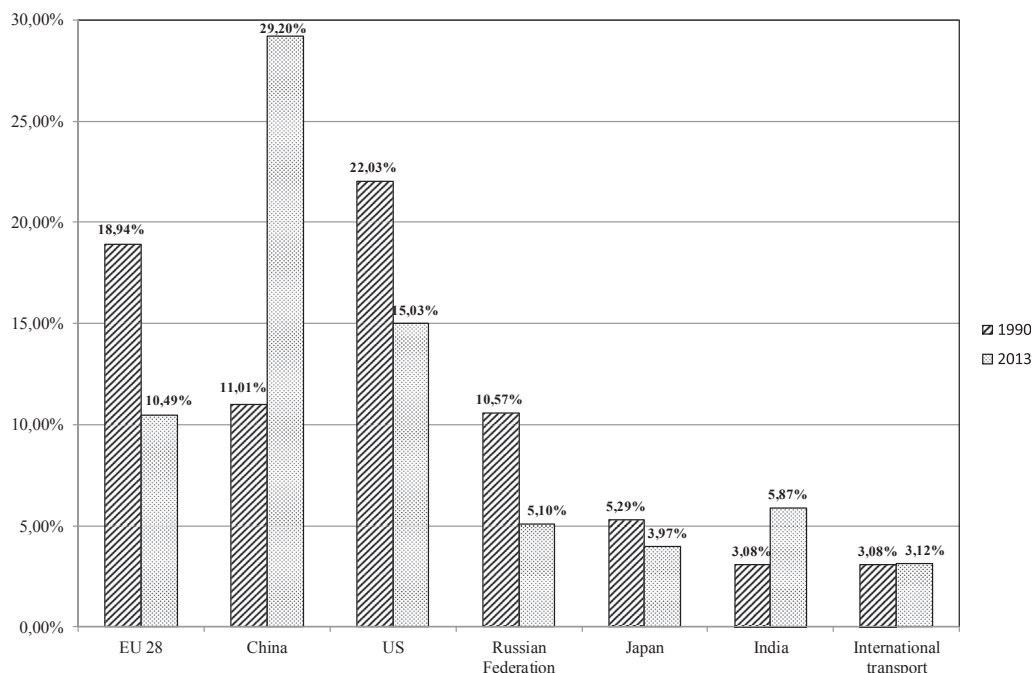


Fig. 2. Share of countries in global emissions of CO₂ in 1990 and 2013.

Source: own work based on (NEAA, 2014).

terms of GHG emission levels. In the analysis the *k-means* method was used, which belongs to the non-hierarchical cluster analysis methods. The study was carried out according to both total emissions, and according to the emissions per capita as well. The analysis was conducted on the basis of Eurostat reports on the GHG emissions of European countries.

2. Research methodology – cluster analysis

The objective of cluster analysis is to group observations of different type in such a way that the resulting clusters are as homogeneous as possible within each group and as different as possible from each other. There is some similarity to the discriminant analysis, where the objective was to obtain a linear combination of the variables such that this transformed linear combination would exhibit the largest difference between centroids but the smallest variance within groups. However, in discriminant analysis, the groups are known a priori, whereas the purpose of cluster analysis is to form such groups; they are called “clusters” (Gatignon, 2010).

There are many methods to create such clusters. The most generally these methods can be divided into hierarchical and non-hierarchical. Among the hierarchical methods agglomerative and divisive methods are distinguished, while among non-hierarchical *k-means* methods, probabilistic clustering and the methods of self-organizing are distinguished. The hierarchical methods allow to create a tree structure (dendrogram) either by treating each element as a separate cluster and merge them in successively larger clusters – in the agglomerative method, or by beginning with the whole set and proceed to divide it into successively smaller clusters – in the divisive method (Madhulatha, 2012).

In the first step of the agglomerative method it is assumed that n observations constitute n separate clusters. On the basis of previously adopted distance measure, e.g. Euclidean distance, the two “nearest” clusters are selected and combined into one single cluster. Thus, one less cluster is obtained, that is $n-1$. The process is repeated until all clusters will be merged into one. It should be

noted that the methods of selecting the nearest (most similar) clusters are few. The most popular are single linkage, complete linkage, centroid method or Ward’s method. Each of these methods can give different results.

The algorithm for divisive method is reverse. At the beginning all the observations are treated as a single cluster, then the following divisions are carried out so long until n clusters are received.

The disadvantage of hierarchical methods is the difficulty to determine the appropriate number of clusters, although, there are some proposals to resolve this problem (Stanisz, 2007; Herbin et al., 2001; Tibshirani et al., 2001).

Among the non-hierarchical methods probably the most popular is the *k-means* method. The name of this method is derived from the representation of each cluster using the average or weighted average. In this method the number of clusters are presumes a priori as well as the number of iterations. Algorithm for creating clusters is strongly dependent on the value of k . The number of clusters should be large enough that clusters will reflect the specific characteristics of the data set. At the same time, however, the value of k must be significantly less than the number of objects in the data set, because that is the meaning of the grouping. In most publications, in which *k-means* method is discussed usually there are suggestions how to determine the appropriate values of k (Pham et al., 2005; Ming-Tso Chiang and Mirkin, 2010).

It should be noted, however, that there is no perfect method for determining k . Often, it is proposed to simply carry out several cycles of calculations for different values of k and based on the obtained statistics (e.g. the inner and the outer distance) to select the most optimal number of clusters. The choice of the best solution is based on a comparison of criterion of internal consistency, based on minimizing deviations within the group (Jain, 2010). Another method – elbow criterion – looks for the percentage of variance explained as a function of the number of clusters. The elbow criterion says that one should choose a number of clusters so that adding another cluster doesn’t add sufficient information. More precisely, if one plots the percentage of variance explained by the

clusters against the number of clusters, the first clusters will add much information, but at some point the marginal gain will drop, giving an angle in the graph (Madhulatha, 2012). Very simple rule of thumb sets the number to $k \cong \sqrt{n/2}$ with n as the number of objects (Mardia et al., 1979). It is also proposed to carry out earlier grouping by hierarchical clustering in order to find the number of clusters, and only then for that value to conduct clustering by *k-means* method (Stanisz, 2007). In further analysis to estimate the number of k the last two methods are used.

In the *k-means* method it is assumed (Morzy, 2013; Stanisz, 2007; Mirkin, 2011) that a set of n objects $D = \{p_1, p_2, \dots, p_n\}$ is given. Each object $p_i = \{p_{i1}, p_{i2}, \dots, p_{im}\}$ represents a point of space R^m , where m is the number of attributes that describe the objects of set D (the number of dimensions of space R^m). Let k is a pre-determined number of clusters, and m_k – the mean of cluster C_k . In the Euclidean space the mean of cluster is calculated according to formula:

$$m_k = \frac{1}{|C_k|} \sum_{p_i \in C_k} p_i$$

In the *k-means* method the aim is to find the division of a set of objects D between the k clusters C_1, C_2, \dots, C_k , of means that minimizes the criterion function $e(k)$ (Likas et al., 2003). In the basic version of the algorithm the criterion function that is minimized, is the sum of the squared error (Larose, 2005):

$$e(k) = \sum_{i=1}^k \sum_{p_j \in C_i} \text{dist}(p_j, m_i)^2,$$

where:

- p_j – point in R^m space that represent object p_j ,
- m_i – mean of cluster C_i ,
- $\text{dist}(p_j, m_i)$ – Euclidean distance (norm L_2) between object (point) p_j and mean (centre) m_i of the nearest cluster C_i .

Algorithm can be described as follows (Han and Kamber, 2006; Hastie et al., 2009; Mirkin, 2011):

A data set containing n objects is given and the number of clusters k is assumed.

1. Arbitrarily choose k objects from D as the initial cluster centres;
2. Repeat steps (a) and (b) until there are changes in the allocation of objects to clusters:
 - a) (re)assign each object $p_i \in D$ to the cluster C_i , to which the object is most similar, based on the mean value of the objects p_i in the cluster C_i ;
 - b) update the cluster means, i.e., calculate the mean value of the objects for each cluster.

K-means method has several advantages. It is relatively simple and the algorithm procedure is relatively efficient compared with hierarchical methods. The reasons for the algorithm's popularity are its ease of interpretation, simplicity of implementation, speed of convergence and adaptability to sparse data (Dhillon and Modha, 2001). Another advantage is its flexibility with regard to the accepted measure of distance. In the basic version of the algorithm it is assumed that $\text{dist}(p_j, m_i)$ is the Euclidean distance (L_2 norm) between the object (point) p_j and the mean (centre) m_i of cluster C_i . But it is possible to use also other distance measures (Singh et al., 2013). Selim and Ismail (1984) gave a rigorous proof of convergence of the *k-means* – algorithm in a generalized form. Also, local

optimality of solutions obtained has been investigated, where it was shown that under certain conditions, the *k-means* algorithm may not yield local minimum solutions. In such cases, means of obtaining local minima were presented.

Main disadvantage of *k-means* is the possibility of getting the distorted results when there are outliers in the data – a single outlier can increase the squared error dramatically (Rokach and Maimon, 2010). Then these more typical objects will tend to be classified into very few groups, but the outliers will tend to be put in very small or even single clusters (Giudici, 2003).

There are many solutions to the problem of outliers. The simplest solution is the removal of the objects, which differ significantly from the centres of clusters. To avoid the error of removing the wrong object, during several iterations of the algorithm it is checked, whether a point, classified as an outlier, differs significantly during the iteration of found clusters. However, in many situations removal of outliers is not allowed. This applies, for example, pattern recognition, image compression, the analysis of financial data, or grouping objects into spatial databases where each points (objects) is subjected to grouping (Morzy, 2013). In the case of grouping the EU countries, each country should be assigned to a cluster. Thus, in presented analysis outliers are not searched.

3. Research results and discussion

3.1. Grouping by *k-means* method according to the general emission levels

In the first stage of the study grouping of the objects – European countries – was conducted according to general levels of emissions of four GHG. These include carbon dioxide (CO_2), nitrogen oxides (NO), methane (CH_4) and nitrous oxide (N_2O). Data were obtained from Eurostat (2015a), which reports data for those gases, although it should be noted that the GHG also includes four other types of gases collectively known as F-gases. Data of GHG emission used for the analysis apply to 25 European Union countries (excluding Malta, Cyprus and Luxembourg), and additionally Norway, Switzerland and Turkey, i.e. 28 countries altogether.

In order to achieve comparability between data collected their normalization method of standardization was made according to the formula (Nisbet et al., 2009):

$$z_i = \frac{x_i - \bar{x}}{S_x}$$

where: \bar{x} and S_x are respectively mean and standard deviation of the variable in experiment.

GHG emissions of EU countries for the year 2012 and their standardized values are presented in Tables 1 and 2.

Using STATISTICA package the groupings of countries was carried out by *k-means* algorithm. In this method, it is first necessary to determine the number of clusters.

In the presented analysis in order to evaluate the values of k , the preliminary calculations have been made with agglomerative method, which resulted in evaluation the number of clusters $k = 4$. To confirm this value additional calculations were carried out with *k-means* method, assuming successively $k = 3, 4$ and 5 . Finally, on the basis of the statistics k was specified as equal 4 . Some confirmation of appropriateness of such a number of clusters is the use of a rule of thumb, whereby $k \cong \sqrt{28/2} \cong 4$.

The initial cluster centres were defined by sorting distances and taking observations at a fixed interval. The resulting grouping and corresponding distances from the cluster centres are shown in Table 3, and the analysis of variance and descriptive statistics in Tables 4 and 5. Cluster no. 1 has 6 countries, including Poland, cluster no. 2 has 2

Table 1
Greenhouse gas emissions by the European countries 2012 [t].

	Carbon dioxide	Nitrogen oxides	Methane	Nitrous oxide
Austria	65 481 220	167 734	252 558	16 385
Belgium	102 316 098	203 640	304 233	22 464
Bulgaria	48 363 949	123 208	342 161	16 219
Croatia	19 422 629	59 002	163 154	10 582
Czech Republic	98 343 688	210 581	476 468	24 527
Denmark	78 117 285	1 089 108	262 535	21 557
Estonia	17 390 938	40 255	44 328	3259
Finland	53 873 231	188 237	194 242	16 031
France	360 238 250	1 058 080	2 401 841	184 352
Germany	890 328 607	2 040 910	2 322 716	181 874
Greece	88 981 101	256 752	463 073	21 325
Hungary	48 627 609	142 979	380 944	21 942
Ireland	38 052 133	74 669	574 956	23 927
Italy	395 866 681	1 041 005	1 655 672	89 729
Latvia	8 764 472	44 428	77 702	5894
Lithuania	18 154 743	81 999	145 796	13 487
Netherlands	203 179 743	485 949	717 835	24 761
Norway	55 211 146	268 391	201 965	10 616
Poland	328 462 794	873 038	1 954 677	95 755
Portugal	51 796 100	186 206	586 721	14 368
Romania	86 651 210	226 835	1 057 039	37 438
Slovakia	35 233 111	80 898	199 105	9019
Slovenia	16 520 613	56 296	88 944	3583
Spain	281 945 961	972 935	1 539 170	59 264
Sweden	53 634 514	255 998	228 030	20 436
Switzerland	47 399 192	91 204	175 703	9836
Turkey	357 498 161	1 088 062	2 934 440	47 700
United Kingdom	536 822 366	1 371 925	2 401 756	115 373

Source: EUROSTAT (2015a).

Table 2
Standardized data – greenhouse gas emissions by the European countries 2012.

	Carbon dioxide	Nitrogen oxides	Methane	Nitrous oxide
Austria	-0.45074	-0.55737	-0.61858	-0.48034
Belgium	-0.26866	-0.48805	-0.55921	-0.35700
Bulgaria	-0.53535	-0.64333	-0.51564	-0.48371
Croatia	-0.67841	-0.76728	-0.72129	-0.59808
Czech Republic	-0.28830	-0.47465	-0.36134	-0.31516
Denmark	-0.38828	1.22141	-0.60712	-0.37540
Estonia	-0.68845	-0.80347	-0.85780	-0.74664
Finland	-0.50811	-0.51778	-0.68558	-0.48752
France	1.00626	1.16150	1.85063	2.92740
Germany	3.62652	3.05892	1.75972	2.87713
Greece	-0.33457	-0.38551	-0.37673	-0.38012
Hungary	-0.53404	-0.60516	-0.47108	-0.36759
Ireland	-0.58632	-0.73703	-0.24819	-0.32733
Italy	1.18237	1.12854	0.99339	1.00767
Latvia	-0.73109	-0.79542	-0.81946	-0.69318
Lithuania	-0.68467	-0.72288	-0.74123	-0.53913
Netherlands	0.22991	0.05697	-0.08405	-0.31040
Norway	-0.50150	-0.36304	-0.67670	-0.59739
Poland	0.84919	0.80427	1.33690	1.12993
Portugal	-0.51838	-0.52170	-0.23468	-0.52126
Romania	-0.34609	-0.44327	0.30565	-0.05321
Slovakia	-0.60025	-0.72501	-0.67999	-0.62979
Slovenia	-0.69275	-0.77250	-0.80655	-0.74008
Spain	0.61926	0.99713	0.85954	0.38960
Sweden	-0.50929	-0.38697	-0.64676	-0.39816
Switzerland	-0.54012	-0.70511	-0.70688	-0.61320
Turkey	0.99272	1.21939	2.46250	0.15499
United Kingdom	1.87913	1.76740	1.85053	1.52795

countries (Denmark and Spain), the third cluster consists of seven objects, while cluster no. 4 – the most numerous – includes 13 countries.

In Table 4 there are measures of intergroup and intragroup diversity of the corresponding variables together with degrees of freedom (*df*). Obtained – as the ratio of intergroup diversity to intragroup diversity – the value of *F* statistics allows for prioritization of variables because of the discriminatory power (Panek,

2009). The analysis of variance shows that the biggest role in the division to clusters by minimizing variability within clusters and maximizing variability between clusters was methane because the value of *F* statistic is the greatest and amounts to 65.47. The second position has nitrogen oxides (*F* = 46.64). On the third and fourth places are respectively, nitrous oxide and carbon dioxide.

On the basis of descriptive statistics contained in Table 5 and the averages presented in Fig. 3 it is seen that 13 countries creating the cluster no. 4 is characterized by the lowest values of average emissions of greenhouse gases. On the other side, cluster no. 1 has the highest GHG emissions. Generally speaking, in this case the smaller the number of cluster, the higher the average values of emitted gases. An exception is cluster no. 2, in which the average standardized values of nitrogen oxides emissions are significantly higher than in clusters no. 3 and no. 4, though smaller than in cluster no. 1.

The most common elements in the fourth cluster are such countries as Switzerland, Slovakia, Lithuania and Croatia because their distances from the centres of clusters are the smallest. Most outstanding countries in this cluster are Sweden, Estonia and Hungary, as evidenced by the greatest distance from the cluster centres. Cluster no. 1 form the major emitters of greenhouse gases in Europe, which includes the United Kingdom, Italy, Poland, France, Turkey and Germany. In this cluster the most typical elements are countries like the United Kingdom, Italy and Poland, while Germany significantly differs from them; its distance from the centre of cluster is almost three times higher than Poland. Table 6 presents information about the Euclidean distances between the centres of clusters and their squares.

Table 7 shows the average values of emitted gases in tones in each cluster.

3.2. Results of the clustering by *k*-means method according to the levels of emissions per capita of European countries

In the second stage of the research the grouping of objects – the European countries – was conducted according to emissions per

Table 3
Elements of clusters with distances form centres for the total emissions.

Cluster 1	Distances from centre of cluster 1	Cluster 2	Distances from centre of cluster 2	Cluster 3	Distances from centre of cluster 3	Cluster 4	Distances from centre of cluster 4
United Kingdom	0.205782	Denmark	0.487452	Czech Republic	0.073653	Switzerland	0.046434
Italy	0.545287	Spain	0.487452	Greece	0.086306	Slovakia	0.051843
Poland	0.597547			Portugal	0.154134	Lithuania	0.069080
France	0.748653			Belgium	0.172588	Croatia	0.079667
Turkey	0.882661			Ireland	0.210582	Finland	0.084685
Germany	1.425747			Romania	0.297642	Austria	0.098621
				Netherlands	0.366363	Bulgaria	0.099560
						Slovenia	0.133499
						Latvia	0.138184
						Norway	0.147746
						Hungary	0.151264
						Estonia	0.155264
						Sweden	0.160044

Table 4
Analysis of variance for total emissions.

Variables	Between SS	df	Inside SS	df	F	Significance p
Carbon dioxide	20.32800	3	6.671997	24	24.37411	0.000000183530800
Nitrogen oxides	23.04719	3	3.952806	24	46.64472	0.000000000363261
Methane	24.06017	3	2.939833	24	65.47357	0.00000000010611
Nitrous oxide	20.35669	3	6.643307	24	24.51393	0.000000174391900

Table 5
Descriptive statistics of clusters on the basis of standardized data for the total emissions.

	Means	Standard deviation	Variance
Cluster 1–6 objects			
Carbon dioxide	1.589366	1.062400	1.128693
Nitrogen oxides	1.523338	0.814103	0.662765
Methane	1.708944	0.502192	0.252196
Nitrous oxide	1.604181	1.100729	1.211604
Cluster 2–2 objects			
Carbon dioxide	0.115492	0.712434	0.507563
Nitrogen oxides	1.109267	0.158590	0.025151
Methane	0.126214	1.037087	1.075549
Nitrous oxide	0.007098	0.540936	0.292611
Cluster 3–7 objects			
Carbon dioxide	-0.301773	0.262962	0.069149
Nitrogen oxides	-0.427606	0.240588	0.057883
Methane	-0.222650	0.275487	0.075893
Nitrous oxide	-0.323496	0.139493	0.019458
Cluster 4–13 objects			
Carbon dioxide	-0.588829	0.094018	0.008839
Nitrogen oxides	-0.643486	0.149035	0.022211
Methane	-0.688273	0.111035	0.012329
Nitrous oxide	-0.567293	0.121080	0.014660

capita of four types of gases (CO₂, NO, CH₄, N₂O). The analysis also applies to the same countries as in Section 3.1. Tables 8 and 9 show GHG emissions per capita of EU countries and their standardized values respectively.

As in Section 3.1 this part of the calculation using firstly agglomerative method and then the *k*-means method for *k* = 3, 4 and 5 the value of *k* was determined as equal 4.

Classification of countries in terms of total emissions of GHG, and emissions per capita show significant differences in obtained clusters. Clusters and distances from the centres of clusters are presented in Table 10. Cluster no. 1 comprises of 9 countries (including Poland), cluster no. 2 contains 6 countries, the third cluster consists of 11 countries, and the fourth cluster has only two

countries (Denmark and Ireland). In Tables 11 and 12 the analysis of variance and descriptive statistics are shown, and in Table 13 information about Euclidean distances between cluster centres and their squares are shown.

The analysis of variance shows that the biggest role in the creating clusters had a variable of nitrous oxide, since the value of *F* is the highest and is equal 22.63. An equally large role variable of CO₂ had played – *F* = 21.99. Less important were emissions (per capita) of methane and nitrogen oxides and these gases had the greatest influence on creating cluster no. 4 (Denmark and Ireland). From Table 12, and Fig. 4 it is seen that in the cluster no. 4 emissions per capita of GHG are the highest, although the level of emissions per capita of CO₂ is close to the level of CO₂ emissions in the cluster no. 1. In clusters no 1, 2 and 3 significant differences exist in emissions per capita of CO₂ and N₂O. The lowest level of CO₂ and CH₄ emissions per capita has reached the countries belonging to the cluster no. 2. Countries in cluster no. 2 are characterized by higher values of nitrous oxide than in cluster no. 1 and no. 3.

The most similar to each other in this respect are such countries as Croatia, France, Latvia and Hungary, because their distances from the centre of cluster are the smallest; while Sweden and Lithuania are countries most outlying in this cluster, because their distances are greatest.

As in Table 12 and Fig. 4 presented means refer to the standardized value, in Table 14 the average emissions of GHG in tones per capita are shown.

Poland has been classified in the group of nine countries in cluster 1, which is in the second place in terms of emissions per capita of CO₂, nitrogen oxides, and methane and in the third place in terms of nitrous oxide. This group also includes the Czech Republic, Germany, Belgium, UK, Finland, the Netherlands and Estonia. The most typical object of this cluster is the Czech Republic, and the most outlying is Estonia. Poland is most similar to Norway and the Netherlands in terms of greenhouse gas emissions per capita.

To the third cluster the following countries were classified: Slovakia, Spain, Italy, Bulgaria, Greece, Slovenia, Austria, Switzerland, Romania, Portugal and Turkey. These countries represent a group of the average level of emissions per capita in comparison with other clusters, i.e. in terms of CO₂ they rank third, in terms of NO and N₂O fourth place, in terms of methane third place. The most typical object of this cluster is Slovakia, and the most outlying is Turkey.

The fourth cluster consists of the major emitters per capita of greenhouse gases in the European countries; it is Ireland and Denmark. The level of average emissions per capita in these two countries in terms of CO₂ is almost twice higher than in the 17

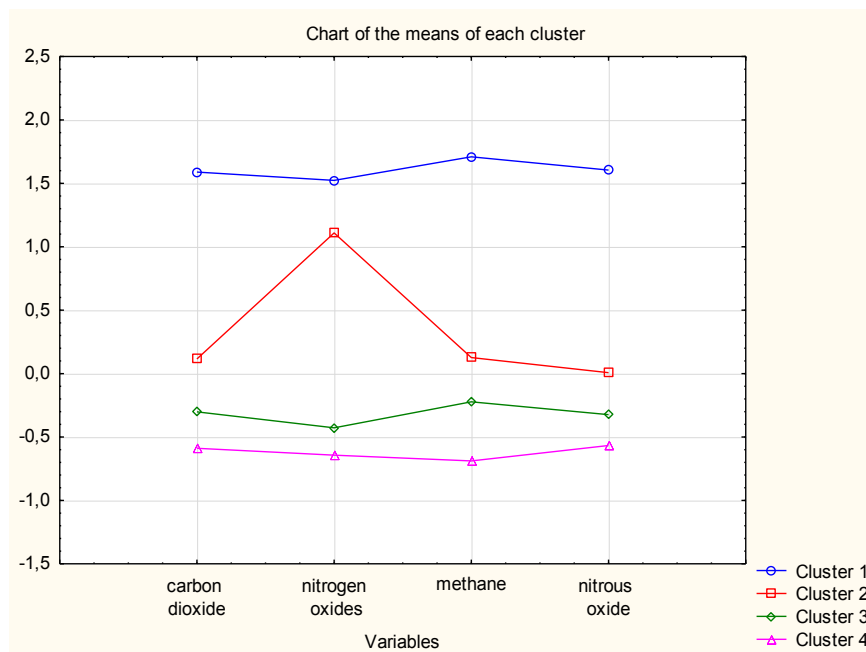


Fig. 3. The mean standardized values in clusters for total emissions.

Table 6

Euclidean distances of clusters (under the diagonal) and squares of distance (above the diagonal).

Number of cluster	No 1	No 2	No 3	No 4
No 1	0.000000	1.849867	3.707396	4.975401
No 2	1.360098	0.000000	0.691772	1.140381
No 3	1.925460	0.831728	0.000000	0.101312
No 4	2.230561	1.067886	0.318295	0.000000

Table 7

The average values of total emissions of gases in clusters [t].

	Carbon dioxide	Nitrogen oxides	Methane	Nitrous oxide
Cluster 1	478 202 810	1 245 503	2 278 517	119 131
Cluster 2	180 031 623	1 031 022	900 852	40 411
Cluster 3	95 617 153	234 948	597 189	24 116
Cluster 4	37 544 413	123 125	191 895	12 099

countries belonging to the second and the third clusters, and slightly higher than in nine countries of the first cluster. The level of emissions of nitrogen oxides is more than three times higher than in other European countries, and the level of methane is over twice as high as all other surveyed countries. The level of nitrous oxide is also almost twice higher than in the other countries.

4. Conclusions and policy implications

The EU countries have seen a steady reduction of GHG emissions, both total and per capita. The total GHG emissions from 1990 to 2012 dropped by 19.2%. To this decline contributed most Germany and the United Kingdom, although both countries had still in 2012 the highest GHG emissions among the EU countries: Germany 20.6%, and UK 13.1%. In the same period also fell in the EU GHG emissions per capita, from the level of 12 tonnes of CO₂ eq. up to 9 tonnes of CO₂ eq. To this decline to the greatest extent contributed the reduction of CO₂ (81.6%), and to a lesser extent, CH₄ (8.9%) and N₂O (7.5%) (EEA, 2014).

Table 8

Greenhouse gas emissions per capita in the European countries 2012 [t/capita].

	Carbon dioxide	Nitrogen oxides	Methane	Nitrous oxide
Austria	7 889 304	20 209	30 429	1974
Belgium	9 744 390	19 394	28 975	2139
Bulgaria	6 363 678	16 212	45 021	2134
Croatia	4 414 234	13 410	37 080	2405
Czech Republic	9 547 931	20 445	46 259	2381
Denmark	14 466 164	201 687	48 618	3992
Estonia	13 377 645	30 965	34 099	2507
Finland	10 164 761	35 516	36 649	3025
France	5 838 545	17 149	38 928	2988
Germany	10 818 088	24 798	28 223	2210
Greece	8 016 315	23 131	41 718	1921
Hungary	4 862 761	14 298	38 094	2194
Ireland	8 648 212	16 970	130 672	5438
Italy	6 675 661	17 555	27 920	1513
Latvia	3 810 640	19 316	33 784	2563
Lithuania	5 339 630	24 117	42 881	3967
Netherlands	12 313 924	29 451	43 505	1501
Norway	11 747 052	57 105	42 971	2259
Poland	8 621 071	22 914	51 304	2513
Portugal	4 886 425	17 567	55 351	1355
Romania	4 030 289	10 550	49 165	1741
Slovakia	6 524 650	14 981	36 871	1670
Slovenia	8 260 307	28 148	44 472	1791
Spain	6 393 332	22 062	34 902	1344
Sweden	5 829 838	27 826	24 786	2221
Switzerland	6 319 892	12 161	23 427	1312
Turkey	4 766 642	14 507	39 126	636
United Kingdom	8 785 963	22 454	39 309	1888

Source: EUROSTAT (2015a).

There are different causes of this decline in GHG emissions in individual countries, as well as different is the level of emissions from various types of gases. Also, various sectors in varying degrees, contribute to GHG emissions. The largest share of these emissions had in 2012 'energy excluding transport' (57.9%), and 'transport including international aviation' (21.9%). The latter rose from 15% in 1990. Lesser share had 'agriculture' (10.0%), 'industrial processes' (6.8%) and 'waste' (3.0%) (EUROSTAT, 2015b).

Table 9
Standardized data – greenhouse gas emissions per capita in the European countries 2012.

	Carbon dioxide	Nitrogen oxides	Methane	Nitrous oxide
Austria	0.03013	-0.23262	-0.60179	-0.30889
Belgium	0.67064	-0.25579	-0.67776	-0.13673
Bulgaria	-0.49663	-0.34629	0.16056	-0.14235
Croatia	-1.16972	-0.42597	-0.25428	0.13957
Czech Republic	0.60281	-0.22591	0.22523	0.11494
Denmark	2.30094	4.92799	0.34845	1.79185
Estonia	1.92511	0.07326	-0.41006	0.24579
Finland	0.81578	0.20267	-0.27680	0.78479
France	-0.67795	-0.31964	-0.15778	0.74645
Germany	1.04136	-0.10211	-0.71705	-0.06343
Greece	0.07398	-0.14953	-0.01199	-0.36400
Hungary	-1.01486	-0.40071	-0.20131	-0.07972
Ireland	0.29216	-0.32472	4.63521	3.29690
Italy	-0.38891	-0.30809	-0.73284	-0.78875
Latvia	-1.37813	-0.25800	-0.42652	0.30380
Lithuania	-0.85021	-0.12148	0.04877	1.76561
Netherlands	1.55783	0.03021	0.08136	-0.80170
Norway	1.36211	0.81656	0.05347	-0.01265
Poland	0.28279	-0.15569	0.48879	0.25236
Portugal	-1.00669	-0.30776	0.70023	-0.95287
Romania	-1.30229	-0.50728	0.37703	-0.55122
Slovakia	-0.44105	-0.38128	-0.26521	-0.62532
Slovenia	0.15822	-0.00685	0.13187	-0.49915
Spain	-0.48639	-0.17993	-0.36810	-0.96497
Sweden	-0.68095	-0.01602	-0.89659	-0.05160
Switzerland	-0.51175	-0.46149	-0.96758	-0.99865
Turkey	-1.04805	-0.39475	-0.14742	-1.70184
United Kingdom	0.33972	-0.16879	-0.13788	-0.39823

Eurostat publishes statistics for the four main types of gases which in this study were included in the analysis. This makes it possible to indicate, for example, the largest and smallest issuers. But it can also be useful to get information about homogeneous groups of countries, taking into account all the analysed types of gases.

To group objects taxonomic method referred to as cluster analysis is preferred. There are many useful methods, hierarchical and non-hierarchical (see Section 2). It should be noted that there may be discussion of which cluster analysis method to apply. As previously mentioned, there is no perfect method, just as there is no ideal method of setting the number of clusters in k-means method.

The aim of the analysis was to examine the diversity of European countries in terms of emissions of four greenhouse gases, i.e. carbon dioxide, nitrogen oxides, methane and nitrous oxide according to total emissions and emissions per capita using one of the non-hierarchical clustering method, i.e. k-means method.

In both analyses, the surveyed EU countries were grouped in four clusters in which the belonging countries are the most

Table 10
Elements of clusters with distances from centres for the emissions per capita.

Cluster 1	Distances from centre of cluster 1	Cluster 2	Distances from centre of cluster 2	Cluster 3	Distances from centre of cluster 3	Cluster 4	Distances from centre of cluster 4
Czech Republic	0.292736	Croatia	0.215059	Slovakia	0.086555	Denmark	1.807408
Germany	0.294114	France	0.215199	Spain	0.172888	Ireland	1.807408
Belgium	0.336748	Latvia	0.231063	Italy	0.294810		
United Kingdom	0.378676	Hungary	0.291224	Bulgaria	0.329576		
Finland	0.413963	Sweden	0.432579	Greece	0.349780		
Norway	0.457261	Lithuania	0.678183	Slovenia	0.399867		
Poland	0.489919			Austria	0.400913		
Netherlands	0.514223			Switzerland	0.436811		
Estonia	0.517335			Romania	0.503036		
				Portugal	0.513332		
				Turkey	0.566979		

Table 11
Analysis of variance for emissions per capita.

	Between SS	df	Inside SS	df	F	Significance p
Carbon dioxide	19.79863	3	7.20137	24	21.99428	0.000000
Nitrogen oxides	11.97196	3	15.02804	24	6.37313	0.002482
Methane	13.49168	3	13.50832	24	7.99014	0.000727
Nitrous oxide	19.94771	3	7.05229	24	22.62837	0.000000

Table 12
Descriptive statistics of clusters on the basis of standardized data for the emissions per capita.

	Mean	Standard deviation	Variance
Cluster 1–9 objects			
Carbon dioxide	0.955349	0.562631	0.316553
Nitrogen oxides	0.023824	0.333287	0.111080
Methane	-0.152300	0.407379	0.165957
Nitrous oxide	-0.001651	0.444041	0.197172
Cluster 2–6 objects			
Carbon dioxide	-0.961969	0.279742	0.078255
Nitrogen oxides	-0.256970	0.160976	0.025913
Methane	-0.314618	0.323832	0.104867
Nitrous oxide	0.470683	0.702296	0.493219
Cluster 3–11 objects			
Carbon dioxide	-0.492675	0.475401	0.226006
Nitrogen oxides	-0.297807	0.146415	0.021437
Methane	-0.156841	0.496808	0.246818
Nitrous oxide	-0.717999	0.433154	0.187622
Cluster 4–2 objects			
Carbon dioxide	1.296551	1.420425	2.01761
Nitrogen oxides	2.301637	3.714224	13.79546
Methane	2.491829	3.031196	9.18815
Nitrous oxide	2.544375	1.064230	1.13259

Table 13
Euclidean distances of clusters (under the diagonal) and squares of distance (above the diagonal) for emissions per capita.

Number of cluster	No 1	No 2	No 3	No 4
No 1	0.000000	1.001100	0.678348	4.694629
No 2	1.000550	0.000000	0.414941	5.955931
No 3	0.823619	0.644159	0.000000	6.904243
No 4	2.166709	2.440478	2.627593	0.000000

homogeneous internally and the most different from each other. There are significant differences in the results of clustering the countries in these two perspectives. Thus using the resulting grouping, it should be decided whether we follow total emission of each country or more important is emission per capita.

The research does not cover wide range of subject of measurement and evaluation of greenhouse gas emissions. It should be

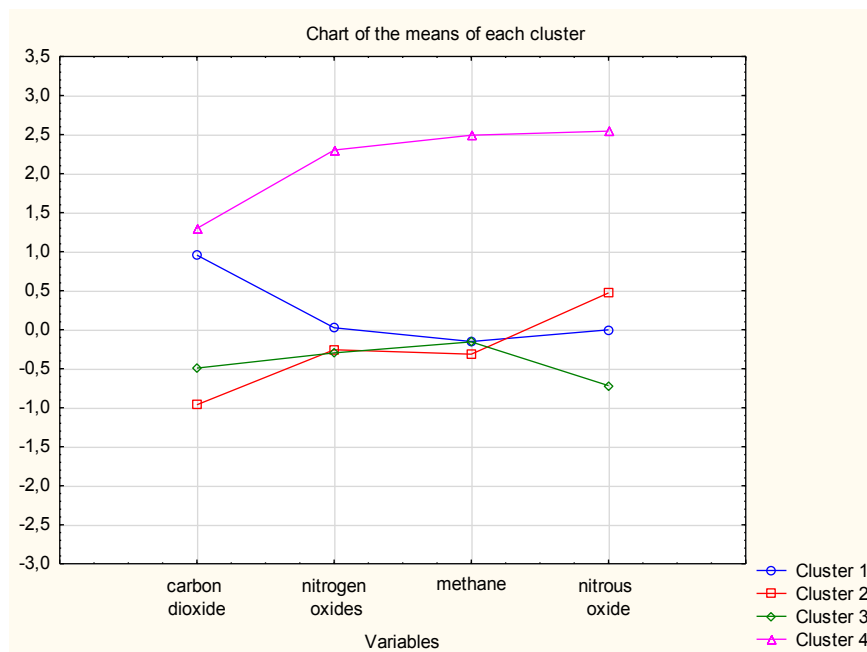


Fig. 4. Chart for the mean values standardized in clusters (emissions per capita).

Table 14

The average values of emissions of gases per capita in clusters [t/capita].

	Carbon dioxide	Nitrogen oxides	Methane	Nitrous oxide
Cluster 1	10 568 980	29 227	39 033	2269
Cluster 2	5 015 941	19 353	35 926	2723
Cluster 3	6 375 136	17 917	38 946	1581
Cluster 4	11 557 188	109 329	89 645	4715

noted that the problem is complex and multi-threaded, among other things, due to the ambiguous impact of emissions on global warming, but also due to the fact that Europe is not the biggest emitter of greenhouse gases in the world. Considering only three major regions of CO₂ emitters, which were responsible in 2013 for more than half of total CO₂ emissions (55%), the European Union emitted just 11% (3.7 billion tonnes), whereas China was responsible for 29% (10.3 billion tonnes of CO₂), and USA for 15% (5.3 billion tonnes CO₂) (NEAA, 2014). Given the analysis presented in the article, it is advisable while setting limits on greenhouse gas emissions to take into account on the one hand, membership of a given country to a particular cluster, and on the other hand to determine the criterion regarding emission limits – total GHG emissions or GHG emissions per capita.

As indicated in the Section 1, scientists are trying to identify the various factors that influence the size and diversity of GHG emissions, as evidenced by, among others, cited earlier publications, however mainly for the China region. Undoubtedly, it would be worth in further studies, based on the experience of Chinese scientists, verify them for the area of Europe.

References

- Burchart-Korol, D., Krawczyk, P., Czaplicka-Kolarz, K., Turek, M., Borkowski, W., 2014. Development of sustainability assessment method of coal mines. *J. Sustain. Min.* 13 (4), 5–11. <http://dx.doi.org/10.7424/jsm140402>.
- Chiang, M.M.-T., Mirkin, B., 2010. Intelligent choice of the number of clusters in K-Means clustering: an experimental study with different cluster spreads. *J. Classif.* 27 (1), 3–40. <http://dx.doi.org/10.1007/s00357-010-9049-5>.

- Dhillon, I., Modha, D., 2001. Concept decompositions for large sparse text data using clustering. *Mach. Learn.* 42 (1–2), 143–175. <http://dx.doi.org/10.1023/A:1007612920971>.
- Dubiński, J., 2013. Sustainable development of mining of mineral resources. *J. Sustain. Min.* 12 (1), 1–6. <http://dx.medra.org/10.7424/jsm130102>.
- Dubiński, J., Turek, M., 2014. Chances and threats of hard coal mining development in Poland – the results of experts research. *Archives Min. Sci.* 59 (2), 395–411. <http://dx.doi.org/10.2478/amsc-2014-0028>.
- EEA (European Environment Agency), 2014. Why Did Greenhouse Gas Emissions Decrease in the EU between 1990 and 2012? <http://www.eea.europa.eu/publications/why-are-greenhouse-gases-decreasing>, accessed in May 2015.
- EO (Energy Outlook), 2015. Energy Outlook 2035 Report accessed in April 2015. <http://www.bp.com/en/global/corporate/about-bp/energy-economics/energy-outlook.html>.
- EUROSTAT, 2015a. Greenhouse gas emissions. Eurostat. http://ec.europa.eu/eurostat/web/products-datasets/-/env_air_gge, accessed in April 2015.
- EUROSTAT, 2015b. Greenhouse Gas Emission Statistics. http://ec.europa.eu/eurostat/statistics-explained/index.php/Greenhouse_gas_emission_statistics#Main_statistical_findings, accessed in September 2015.
- Gatignon, H., 2010. *Statistical Analysis of Management Data*. Springer.
- Giudici, P., 2003. *Applied Data Mining. Statistical Methods for Business and Industry*. John Wiley & Sons.
- GUS, 2014. *Central Statistical Office, Energy*. Warsaw 2014.
- Hąbek, P., Wolniak, R., 2016. Assessing the quality of corporate social responsibility reports; the case of reporting practices in selected European Union member states. *Springer Qual. Quant.* 50 (1), 399–420.
- Hąbek, P., 2012. Sustainability report. Disclosure the impact of business on society and the environment. *J. Marit. Univ. Szczecin* 31 (103), 83–88. <http://repository.scientific-journals.eu/handle/123456789/387>.
- Hąbek, P., 2014. Evaluation of sustainability reporting practices in Poland. *Springer Qual. Quant.* 48, 1739–1752. <http://dx.doi.org/10.1007/s11135-013-9871-z>.
- Han, J., Kamber, M., 2006. *Data Mining: Concepts and Techniques*, second ed. Elsevier MK Morgan Kaufmann Publishers.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, second ed. Springer.
- Herbin, M., Bonnet, N., Vautrot, P., 2001. Estimation of the number of clusters and influence zones. *Pattern Recognit. Lett.* 22 (14), 1557–1568. [http://dx.doi.org/10.1016/S0167-8655\(01\)00103-9](http://dx.doi.org/10.1016/S0167-8655(01)00103-9).
- Idso, C.D., Carter, R.M., Singer, S.F., Soon, W., 2013. Scientific Critique of IPCC's, Summary for Policymakers. Centre for the study of carbon dioxide and global change, The Heartland Institute, Science and Environmental Policy Project. https://www.heartland.org/sites/default/files/critique_of_ipcc_spm.pdf, accessed in April 2015.
- IEA, 2014. CO₂ emissions from Fuel Combustion. Highlights. 2014 Edition. International Energy Agency. <https://www.iea.org/publications/freepublications/publication/co2-emissions-from-fuel-combustion-highlights-2014.html>, accessed in May 2015.
- Jain, A.K., 2010. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* 31 (8), 651–666. <http://dx.doi.org/10.1016/j.patrec.2009.09.011>.

- Krzemień, A., Więckol–Ryk, A., Duda, A., Koterak, A., 2013. Risk assessment of a post – combustion and amine – based CO₂ capture ready process. *J. Sustain. Min.* 12 (4), 18–23. <http://dx.doi.org/10.7424/jsm130404>.
- Larose, D.T., 2005. *Discovering Knowledge in Data. An Introduction to DATA MINING*. John Wiley & Sons.
- Li, J.S., Xia, X.H., Chen, G.Q., Alsaedi, A., Hayat, T., 2016. Embodied energy based optimal energy abatement strategy for Beijing economy. *Renew. Sustain. Energy Rev.* 53, 1602–1610. <http://dx.doi.org/10.1016/j.rser.2015.09.090>.
- Likas, A., Vlassis, N., Verbeek, J.J., 2003. The global k-means clustering algorithm. *Pattern Recognit.* 36 (2), 451–461. [http://dx.doi.org/10.1016/S0031-3203\(02\)00060-2](http://dx.doi.org/10.1016/S0031-3203(02)00060-2).
- Lutyński, M., 2014. Impact of preparation and storage of activated carbon on the high pressure sorption of CO₂. *Bull. Pol. Acad. Sci. Tech. Sci.* 62 (1), 113–119.
- Madhulatha, T.S., 2012. An overview on clustering methods. *IOSR J. Eng.* 2 (4), 719–725. <http://dx.doi.org/10.9790/3021-0204719725>. Apr.
- Majer, M., 2013. The practice of mining companies in building relationships with local communities in the context of CSR formula. *J. Sustain. Min.* 12 (3), 38–47. <http://dx.doi.org/10.7424/jsm130305>.
- Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. *Multivariate Analysis. Probability and Mathematical Statistics*. Academic Press.
- McKinsey, 2010. *Assessment of the Potential Reduction of Greenhouse Gas Emissions in Poland to 2030*. McKinsey&Company, Warsaw.
- Mirkin, B., 2011. *Core Concepts in Data Analysis: Summarization, Correlation and Visualization*. Springer-Verlag London Ltd.
- Morzy, T., 2013. *Data Exploration. Methods and Algorithms*. PWN Warsaw.
- MŚ, 2003. *Polityka klimatyczna Polski. Strategie Redukcji Emisji gazów Ciężkich w Polsce do Roku 2020*. Ministry of Environment in Poland Warszawa.
- NEAA, 2014. *2014-Trends in Global CO₂ Emissions*. PBL Netherlands Environmental Assessment Agency. Report. <http://www.pbl.nl/en/publications/trends-in-global-co2-emissions-2014-report>. accessed in May 2015.
- Nisbet, R., Elder, J., Miner, G., 2009. *Handbook of Statistical Analysis and Data Mining Applications*. Elsevier Academic Press.
- Olajire, A.A., 2013. A review of mineral carbonation technology in sequestration of CO₂. *J. Pet. Sci. Eng.* 109, 364–392. <http://dx.doi.org/10.1016/j.petrol.2013.03.013>.
- Pachauri, R.K., Meyer, L.A., 2014. *Climate Change 2014. Synthesis Report*. http://www.ipcc.ch/pdf/assessment-report/ar5/syr/SYR_AR5_FINAL_full.pdf. accessed in April 2015.
- Panek, T., 2009. *Statistical Methods of Multidimensional Comparative Analysis*. SGH, Warsaw.
- Pawłowski, A., Cao, Y., 2014. The role of CO₂ in the Earth's ecosystem and the possibility of controlling flows between subsystems. *Miner. Resour. Manag.* 30 (4), 5–20. <http://dx.doi.org/10.2478/gospo-2014-0037>.
- PEP (Polish Energy Policy), 2014. *Polish Energy Policy 2050*. Warsaw 2014.
- Pham, D.T., Dimov, S.S., Nguyen, C.D., 2005. Selection of K in K-means clustering. *Proc. IMechE Part C J. Mech. Eng. Sci.* 219, 103–119. <http://dx.doi.org/10.1243/095440605X8298>.
- Pietrzyk-Sokulska, E., Uberman, R., Kulczycka, J., 2015. The impact of mining on the environment in Poland – myths and reality. *Miner. Resour. Manag.* 31 (1), 45–64. <http://dx.doi.org/10.1515/gospo-2015-0009>.
- Pitfield, P.E.J., Brown, T.J., Idoine, N.E., 2010. Mineral information and statistics for the BRIC countries 1999–2008. *Br. Geol. Surv.* <https://www.bgs.ac.uk/downloads/start.cfm?id=1637> accessed in July 2015. p. 3.
- Ranosz, R., 2014. Mining and its importance in the global economy. *Miner. Resour. Manag.* 30 (1), 5–20. <http://dx.doi.org/10.2478/gospo-2014-0003>.
- Rokach, L., Maimon, O. (Eds.), 2010. *Data Mining and Knowledge Discovery Handbook*, second ed. Springer. <http://dx.doi.org/10.1007/978-0-387-09823-4>.
- Selim, S., Ismail, M., 1984. K-means-type algorithms: a generalized convergence theorem and characterization of local optimality. *IEEE Trans. Pattern Anal. Mach. Intell.* 1, 81–87. <http://dx.doi.org/10.1109/TPAMI.1984.4767478>.
- Singh, A., Yadav, A., Rana, A., 2013. K-means with three different distances metrics. *Int. J. Comput. Appl.* 67 (10), 13–17. <http://dx.doi.org/10.5120/11430-6785>. April.
- Stanisz, A., 2007. *Accessible statistics course using STATISTICA examples from medicine*. *Multivar. Anal.* 3. Krakow.
- Tibshirani, R., Guenther, W., Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. B* 63 (2), 411–423. <http://dx.doi.org/10.1111/1467-9868.00293>.
- Uliasz-Bocheńczyk, A., 2010. Mineral sequestration of CO₂ in suspensions containing mixtures of fly ashes and desulphurization waste. *Miner. Resour. Manag.* 26 (4), 109–118. <http://dx.doi.org/10.2478/gospo-2013-0037>.
- Uliasz-Bocheńczyk, A., Mokrzycki, E., 2014. Mineral carbonation using natural materials – CO₂ reduction method? *Miner. Resour. Manag.* 30 (3), 99–110. <http://dx.doi.org/10.2478/gospo-2014-0027>.
- Uliasz-Bocheńczyk, A., Mokrzycki, E., Piotrowski, Z., Pomykała, R., 2009. Estimation of CO₂ sequestration potential via mineral carbonation in fly ash from lignite combustion in Poland. *Energy Procedia* 1, 4873–4879. <http://dx.doi.org/10.1016/j.egypro.2009.02.316>.
- Xia, X.H., Hu, Y., Alsaedi, A., Hayat, T., Wu, X.D., Chen, G.Q., 2015. Structure decomposition analysis for energy-related GHG emission in Beijing: urban metabolism and hierarchical structure. *Ecol. Inf.* 26, 60–69. <http://dx.doi.org/10.1016/j.ecoinf.2014.09.008>.
- Xia, X.H., Chen, G.Q., 2012. Energy abatement in Chinese industry: cost evaluation of regulation strategies and allocation alternatives. *Energy Policy* 45, 449–458. <http://dx.doi.org/10.1016/j.enpol.2012.02.055>.
- Xia, X.H., Huang, G.T., Chen, G.Q., Bo, Zhang, Chen, Z.M., Yang, Q., 2011. Energy security, efficiency and carbon emission of Chinese industry. *Energy Policy* 39, 3520–3528. <http://dx.doi.org/10.1016/j.enpol.2011.03.051>.