Ecological and economic aspects of sustainable development of Ukrainian regions

Vitalina Babenko¹, Larysa Zomchak², and Maryna Nehrey^{3,}*

¹V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine

²Ivan Franko National University of Lviv, 1 Universytetska Str., Lviv, 71001, Ukraine

³National University of Life and Environmental Sciences of Ukraine, 15 Heroyiv Oborony Str., Kyiv, 03041, Ukraine

Abstract. The need for sustainable development of Ukrainian regions is due to the global threat of environmental degradation, the unstable situation in the world economy, low socio-economic indicators of the country and weak innovation activity. An important factor that has an influence on the economic development of regions is the ecological state of the environment. It plays one of the most important roles in the conducting of economic activities that require the use of natural resources. According to the results of the investigation, four clusters were formed. Cluster analysis made it possible to conduct a general assessment of the state of the regions of Ukraine, to form groups by similarity and to draw sound conclusions about the existence of similarities in the economy. The formation of clusters and the development of sustainable development policies for individual clusters, which should have significant differences, taking into account their specifics, will contribute to the more effective achievement of sustainable development goals.

1 Introduction

Currently, sustainable development is becoming an important component of the development of Ukraine and its regions. However, the unstable economic situation in the country, raw materials, and export-oriented nature of the country's economy do not allow it to effectively use the mechanisms of sustainable development policy developed for developed countries.

Ecological and economic aspects of sustainable development are in fact the separation of production processes and environmental protection. When developing a policy for the sustainable development of regions, it is important to take into account both the socio-economic and environmental interests of the local population. Underestimation of the environmental factor and environmental constraints in the development of sustainable development documents leads not only to numerous negative consequences in nature management, but also deep long-term disparities between economic, social, and environmental development of different levels, affects the quality and efficiency of environmental regulation. In this regard, taking into account the environmental factor in the management of sustainable development of the region becomes especially important.

In crisis and an unstable external environment, the regions of Ukraine are faced with the problem of sustainable and stable socio-economic development. Sustainable regional development in a crisis is of particular importance and significance, since without clear goals and guidelines for movement, the achieved results can be negative. The cluster approach, which allows government bodies to achieve a long-term competitive advantage, can become the basis for

sustainable development of the region. The idea of the investigation is to substantiate a theoretical and practical approach to assessing the need to form the regional potential of cluster formation for sustainable regional development. The purpose of the study is to use data mining tools to identify patterns and relationships based on data on air pollution in the regions of Ukraine. Spatial data mining involves identifying interesting and potentially useful templates from databases by grouping objects into clusters.

2 Sustainable development of Ukrainian regions

The need for sustainable development of Ukrainian regions is due to the global threat of environmental degradation, the unstable situation in the world economy, low socio-economic indicators of the country and weak innovation activity. The basic aspects of sustainable development policy, goal setting and interaction between them, as well as the experience of developed countries are presented in studies [1-7].

Competitive regions of Ukraine are a source of growth for the whole country, a pillar of national policy to reduce regional disparities, promote more balanced development and promote sustainable development of the country. In general, ensuring sustainable development of the country is possible only when sustainable development of all regions is achieved. The policy of sustainable development of the regions provides for the formation of such conditions and the use of such mechanisms under which the natural basis of this development is not

* Corresponding author: marina.nehrey@gmail.com

destroyed, the preservation and reproduction of the environment is ensured.

Sustainable development of the region is an activity of planning, organization, coordination of economic, social, technical-technological, ecological, and other processes, aimed at effective use of its economic potential in order to improve the quality of life and working conditions of economic entities. not only present but also future generations. Moreover, this activity is carried out both within the region and abroad. Analysis of sustainable development of regions, their features and prospects for implementation are presented in [8-17]. Measurement of various aspects (ecological, financial, technical etc.) of regional development and analysis of their impact on sustainable development presented in the papers [18-23].

Sustainable development policy at the regional level should be aimed at achieving the following goals:

- greening of the economy;
- improving the quality of the environment;
- improving the quality of life of the population;
- restoration of natural resources.

The main tools for developing sustainable development policy are:

- at the state level – coordination of infrastructure development, granting special status to certain territories, direct financial support (subsidies, transfers);
- at the regional level – strategic planning of regions in order to balance socio-ecological and economic development, limiting the creation of environmentally harmful enterprises in densely populated areas;
- at the community level – cooperation between communities to join forces to accelerate development.

3 Cluster analysis as method of organizing groups of objects

Cluster analysis is used in various fields and industries. It works even when there is little data and the requirements for the normality of the distribution of random variables and other requirements of classical methods of statistical analysis are not met. It is useful when you need to classify a large amount of information.

Cluster analysis has a number of advantages over other methods of data classification. First, it allows to break down objects by a single feature or by a whole set of features. Moreover, the influence of each of the parameters can be quite simply enhanced or weakened by making the appropriate coefficients in the mathematical formulas. Second, cluster analysis does not restrict the type of grouping objects and allows to consider many data source of almost arbitrary nature. Third, the peculiarity of clustering is that most algorithms are able to independently determine the number of clusters into which you want to break down the data, as well as to identify the characteristics of these clusters without human intervention, only using the algorithm used. The essence of the cluster analysis procedure is that the objects are represented by a vector (set) of individual features of these objects in the form of a table “object-property”, on the basis of which the matrix of distances (similarity, proximity) is calculated, which is carried out by

clustering. This solves the problem of classifying data using a well-formed mathematical apparatus.

Cluster analysis of the spatial distribution of objects allows to reduce the dimensionality of data, to make it clear. Thus, whenever a large amount of information needs to be classified into groups suitable for further processing, cluster analysis is very useful and effective. The main task of cluster analysis is the formation of homogeneous groups in multidimensional space.

Clustering algorithms are usually built as a certain way to search the number of clusters and determine its optimal value in the search process and include 5 basic steps:

1. Sampling for clustering.
2. Determining the criteria by which objects will be evaluated in the sample.
3. Calculation of values of one or another degree of similarity between objects.
4. Application of cluster analysis to create groups of similar objects.
5. Verification of the results of the cluster solution.

Today, there are many methods of dividing groups of objects into clusters. There are several dozen algorithms and even more modifications. But most often the methods of cluster analysis are divided into two large groups: hierarchical and non-hierarchical.

When choosing between hierarchical and non-hierarchical methods, it is necessary to take into account their features. Non-hierarchical methods show higher resistance to noise and emissions, incorrect choice of metrics, the introduction of insignificant variables in the set involved in clustering. The price to pay for these benefits of the method is the word "a priori". The analyst must determine in advance the number of clusters, the number of iterations or the stop rule, as well as some other clustering parameters. This is especially difficult for novice professionals. If there are no assumptions about the number of clusters, it is recommended to use hierarchical algorithms. However, if the sample size does not allow this, a possible way is to conduct a series of experiments with different numbers of clusters, for example, to start breaking down the data set from two groups and, gradually increasing their number, to compare the results. Due to this "variation" of the results, a fairly high flexibility of clustering is achieved. Hierarchical methods, in contrast to non-hierarchical ones, refuse to determine the number of clusters, and build a complete tree of nested clusters. Complexities of hierarchical clustering methods: limiting the scope of the data set, choosing the degree of proximity, inflexibility of the obtained classifications. The advantage of this group of methods compared to non-hierarchical methods is their clarity and the ability to obtain a detailed representation of the data structure.

Thus, the following conclusions can be drawn: cluster analysis is a universal tool that can be used in regional modelling. With its help, you can analyse data on the similarity of objects. The results of the analysis are presented in a convenient visual form, which facilitates decision-making to determine the optimal number of factors and the relationship of various clusters.

4 Regions of Ukraine clustering by ecological and economic indicators

4.1 Regions of Ukraine clusterization by the level of air pollution

An important factor that has an influence on the economic development of regions is the ecological state of the environment. It plays one of the most important roles in the conducting of economic activities that require the use of natural resources. Personnel potential often depends on the ecological condition of the region, people have a desire to live and work in cities with good environmental performance, access to drinking water, and clean air.

For investigating the economic potential of the regions, it is advisable to form a comprehensive assessment of the regional environmental conditions according to several indicators that are most important. To characterize the quality of the ecological state, the regions should be classified according to several diverse quantitative characteristics in order to identify homogeneous and unique objects by the obtained values. The most effective method for such a multicriteria classification is cluster analysis.

The subject of this investigation is the application of clustering methods for the formation of cluster groups from the regions of Ukraine. The comprehensive assessment will be based on the comparison of all regions, the formation of groups on the similarity of their characteristics with each other, using the techniques of multidimensional classification of economic objects to combine them into homogeneous classification groups on selected characteristics.

The aim of the investigation is a clustering of regions of Ukraine by air pollution. Statistical data from the Statistical Yearbook Environment of Ukraine 2018 (published by the State Statistics Service of Ukraine) on air pollution in the regions of Ukraine by emissions of nitrogen dioxide (thousands of tons) and emissions of sulfur dioxide (thousands of tons) are used [24].

Unlike controlled learning, where there is a base of values to assess the effectiveness of the model, the cluster method k-means does not have a solid evaluation metric that we can use to evaluate the results of different clustering algorithms. Moreover, the k-mean method requires an input number k , and this value is not calculated from the values that need to be evaluated. As a result, there is no right answer to what number of clusters we need to build. Sometimes subject knowledge or intuition can help you choose the number k , but this is not always true. According to the methodology of cluster forecasting, we can assess how well the models work based on different numbers of clusters. In this paper, we will consider 2 indicators that can help us estimate how many clusters we need to use in the k-mean method: elbow method and Silhouette analysis.

The Elbow method will help calculate the optimal number of clusters based on the sum of the squares of the distances between the data points and the centroids. When visualizing this method, the optimal value is where the curve is the first bent, and then gradually aligned.

Fig. 1. Calculation of the number of clusters with the Elbow method.

As we can see, the graph from Fig. 1 indicates that the number of clusters should be 2. However, sometimes it is difficult to determine the exact number of clusters, because the curve may fall monotonically and will not be visually visible moment of its “fracture” or “elbow”.

Silhouette analysis can be used to determine the degree of division between clusters. For each example you need:

1. Calculate the average distance from all data points in one cluster (a^i).
2. Calculate the average distance from all data points in the nearest cluster (b^i).
3. Calculate the coefficient:

$$\frac{b^i - a^i}{\max(a^i, b^i)}.$$

The coefficient can take values in the range $[-1; 1]$:

- If it is equal to 0 – then the element is very close to neighboring clusters.
- If it is equal to 1 – then the element is far from neighboring clusters.
- If it is equal to -1 – then the element is improperly classified, belongs to the wrong cluster.

As we see from Fig. 2 – 4, the coefficient from the Silhouette analysis method is closest to 1 when constructing two clusters and is equal to 0.8 (Fig. 2). We obtained good results, which confirm the result of the previous method Elbow method, that the optimal number of clusters for clustering our data by the k-mean method is the number 2. So, both Silhouette analysis method and Elbow method has given the same result for the number of the clusters, that’s why regions of Ukraine were grouped in two clusters.

Clustering results:

1. The first cluster includes Donetsk, Dnipropetrovsk and Ivano-Frankivsk regions.
2. The second cluster includes Vinnytsia, Volyn, Zhytomyr, Zakarpattia, Zaporizhia, Kyiv, Kirovohrad, Luhansk, Lviv, Mykolaiv, Odesa, Poltava, Rivne, Sumy,

Ternopil, Kharkiv, Kherson, Khmelnytsky, Cherkasy, Chernihiv and Chernivtsi, Kyiv.

Fig. 2. Silhouette analysis for two clusters.

Fig. 3. Silhouette analysis for three clusters.

Fig. 4. Silhouette analysis for four clusters.

The centroid for the first cluster contains a pollution value of 237.5 thousand tons of sulfur dioxide emissions and 37.2 thousand tons of nitrogen dioxide emissions into the atmosphere; for the second cluster 16.54 thousand tons and 5.83 thousand tons, respectively.

4.2 Clustering of Ukraine regions by environmental factors

It is obvious that the ecological condition of the regions of Ukraine depends not only on the state of air pollution. Therefore, in addition to the characteristics of air purity, the study includes the number of green areas, the degree of environmental pollution and the capacity of treatment plants.

The criterias that will characterize the state of the environment in all areas and prospects for its improvement:

- emissions of pollutants into the atmosphere from stationary sources of pollution by region (t / km^2);

- area of forests reproduction by regions (ha);
- waste incineration by regions (thousand tons);
- capacity of treatment plants by regions (million m^3).

All statistical indicators are taken from Statistical Yearbook Environment of Ukraine 2018 (by the State Statistics Service of Ukraine) [24].

Four clusters were obtained by the hierarchical clustering method:

- The first cluster. Donetsk and Dnipropetrovsk regions;
- The second cluster. Chernivtsi, Khmelnytsky, Ternopil, Cherkasy, Sumy, Luhansk, Kherson, Mykolaiv, Kirovohrad regions;
- The third cluster. Rivne, Zhytomyr, Chernihiv, Zakarpattia, Kyiv, Volyn;
- The fourth cluster. Ivano-Frankivsk, Kharkiv, Odesa, Zaporizhia, Lviv, Poltava and Vinnytsia regions.

The dendrogram is shown in the figure 5.

1 – Vinnytsya, 2 – Volyn, 3 – Dnipropetrovsk, 4 – Donetsk, 5 – Zhytomyr, 6 – Zakarpattia, 7 – Zaporizhzhya, 8 – Ivano-Frankivsk, 9 – Kyiv, 10 – Kropyvnytskyi, 11 – Luhansk, 12 – Lviv, 13 – Mykolayiv, 14 – Odesa, 15 – Poltava, 16 – Rivne, 17 – Sumy, 18 – Ternopil, 19 – Kharkiv, 20 – Kherson, 21 – Khmelnytskyi, 22 – Cherkasy, 23 – Chernivtsi, 24 – Chernihiv.

Fig. 5. Diagram of hierarchical cluster of regions of Ukraine by environmental conditions.

For the accuracy of the study, we form four clusters using another method – k-means.

We get the result, where for each cluster there is the region and the distance from the cluster center:

- the first cluster: Dnipropetrovsk 0.393, Donetsk 0.393;
- the second cluster: Zakarpattia 0.377, Kropyvnytskyi 0.268, Luhansk 0.344, Mykolayiv 0.378, Sumy 0.103, Ternopil 0.293, Kherson 0.310, Khmelnytsky 0.194, Cherkasy 0.1, Chernivtsi 0.186, Chernihiv 0.544;
- the third cluster: Volyn 0.332, Zhytomyr 0.548, Kyiv 0.544, Rivne 0.462;
- the fourth cluster: Vinnytsia 0.376, Zaporizhzhya 0.538, Ivano-Frankivsk 1.202, Lviv 0.527, Odesa 0.416, Poltava 0.529, Kharkiv 0.653.

The means of the emissions of pollutants into the atmosphere from stationary sources of pollution by cluster for the first cluster is 702,25 for the second cluster is 21,07\$ for the third cluster is 21,13 and for the fourth cluster is 104,9 t / km^2 .

The means of the area of forests reproduction by regions for each cluster respectively are 245,5; 1389,9; 6185,25 and 1537 ha.

The waste incineration by clusters are 15; 17,25; 37,1 and 65,35 (thousand tons);

The capacity of treatment plants by clusters are 792,5; 85,63; 110,2; 250,56 (million m³).

As we can see, the two regions with the largest number of industrial facilities in Donetsk and Dnepropetrovsk, are immediately combined into one group, which can be explained by voluminous emissions of pollutants into the atmosphere from production, a large amount of incinerated waste, and a small area of the forest reproduction, given the location in the steppe zone.

In the second cluster, there are the objects that have the closest connection with each other – these are the regions of the central and partly western part of the country. They have a similar relief, the average man-made load on the environment. They can be described as areas with an ecological situation that needs improvement, due to the planting of forests and increasing the number of treatment facilities. It should be noted that this includes the Zakarpattia region, which does not have a large number of industrial facilities, but there is a problem of mass deforestation, which has grown into a real environmental disaster.

The third group is formed in regions that have a medium close connection with each other. These are the regions of Ukraine where the ecological situation is considered the best. Volyn, Zhytomyr, Rivne regions have the best indicators of air purity, there are no heavy industry enterprises that carry out large-scale emissions of waste. The areas are located in a forest area, which only benefits the ecology of the areas.

The fourth cluster regions, as well as the previous one of medium-density connections between objects. These are the regions of the southern and western parts of Ukraine, without a large number of industrial facilities with emissions of harmful substances. Areas of forest and steppe zones, with treatment facilities of medium capacity.

4.3 Regions cluster analysis taking into account the gross regional product

For investigation the relationship between environmental indicators and economic potential, we will carry out clustering of regions based on gross regional product (figure 6).

Three groups of objects are now clearly distinguished by similarity. The first cluster remained unchanged: Donetsk and Dnipropetrovsk oblasts, while the second was supplemented by Chernihiv, Zakarpattia, and Volyn oblasts as a result of the reduction of distances to the cluster center. The other three areas in the third group became similar to the objects in the next cluster.

For a more detailed distribution we use the method of k-means:

– the first cluster: Volyn 0.822, Transcarpathian 0.822, Kropyvnytskyi 0.288, Luhansk 0,358, Mykolayiv 0.414, Sumy 0.08, Ternopil 0.314, Kherson 0.345,

Khmelnytsky 0.203, Cherkasy 0.219, Chernivtsi 0.352, Chernihiv 0.42;

– the second cluster: Dnipropetrovsk 0.608, Donetsk 0.608;

– the third cluster: Vinnytsia 0.382, Zhytomyr 1.257, Zaporizhzhya 0.696, Ivano-Frankivsk 1.189, Kyiv 0.764, Lviv 0.324, Odessa 0.633, Poltava 0.516, Rivne 0.462, Kharkiv 0.959.

1 – Vinnytsya, 2 – Volyn, 3 – Dnipropetrovsk, 4 – Donetsk, 5 – Zhytomyr, 6 – Zakarpattia, 7 – Zaporizhzhya, 8 – Ivano-Frankivsk, 9 – Kyiv, 10 – Kropyvnytskyi, 11 – Luhansk, 12 – Lviv, 13 – Mykolayiv, 14 – Odesa, 15 – Poltava, 16 – Rivne, 17 – Sumy, 18 – Ternopil, 19 – Kharkiv, 20 – Kherson, 21 – Khmelnytskyi, 22 – Cherkasy, 23 – Chernivtsi, 24 – Chernihiv.

Fig. 6. Diagram of the results of cluster analysis taking into account the gross regional product.

The distance to the centers of clusters has increased for all regions, which is due to the addition of another characteristic of the studied objects. The number of groups also changed, with an increase in the similarity between the objects of each group. This indicates that we have been able to explore in more detail the relationship between environmental performance and economic development. The greater the similarity of the regions with each other in terms of the environment, the greater the similarity between them in economic characteristics.

Cluster analysis of the data made it possible to conduct a general assessment of the state of the regions of Ukraine, to form similarity groups, and to make reasonable conclusions about the existence of similarities in economic development and shortcomings that should be addressed.

5 Conclusions

According to the results of the investigation, four clusters were formed, the first of which included regions with developed heavy industry: Donetsk and Dnipropetrovsk, which characterizes them as regions with the worst environmental status. The second cluster – 12 regions of central and eastern Ukraine with a lower level of pollution than in the previous group, but still not with the best performance. This group includes the objects with the closest connection. The other two clusters have a medium-tight relationship, and are characterized by some of the best environmental performance. Cluster analysis made it possible to conduct a general assessment of the state of the regions of Ukraine, to form groups by

similarity and to draw sound conclusions about the existence of similarities in the economy.

The sustainable development policy design should be implemented taking into account the specifics of the regions. In some regions, the pace and timing of changes and development may be different. The features of sustainable development are due to the fact that the goals and conditions of socio-economic development, the use of natural resources, and environmental protection are inextricably linked with a certain territory, which is characterized by specific geographic, demographic, and economic conditions. The regions of Ukraine differ significantly from each other in the area of territories, population size, volumes of industrial and agricultural production, the average per capita real incomes, which predetermines the differentiation of regions in terms of the available economic potential, level, and populations life quality.

This explains the fact that a universal sustainable development policy for all regions cannot be developed. Therefore, each region needs to design its own sustainable development policy as part of the country's overall sustainable development policy. The formation of clusters and the development of sustainable development policies for individual clusters, which should have significant differences, taking into account their specifics, will contribute to the more effective achievement of sustainable development goals.

References

1. D. Dizdaroglu. Sustainability **9(6)**, 1018 (2017)
2. M. Nilsson, E. Chisholm, et al. Sustainability science **13(6)**, 1489 (2018)
3. E. Holden, K. Linnerud, D. Banister. Sustainable Development **25(3)**, 213 (2017)
4. C. Allen, G. Metternicht, T. Wiedmann. Sustainability Science **13(5)**, 1453 (2018)
5. P. Hryhoruk, N. Khrushch, S. Grygoruk. E3S Web of Conferences **166**, 13023 (2020)
6. O. Kuzmenko, T. Vasylieva, S. Vojtovič, O. Chygryn, V. Snieska. Economics and Sociology **13(4)**, 318 (2020)
7. N. Davydenko, S. Kvasha, Y. Pasichnyk, T. Viatkina, N. Wasilewska. Problems and Perspectives in Management **16(4)**, 186 (2018)
8. R. Jovovic, M. Draskovic, M. Delibasic, M. Jovovic. Journal of International Studies **10(1)** (2017)
9. J. Kono, Y. Ostermeyer, & H. Wallbaum. Journal of Cleaner Production **196**, 1356 (2018)
10. R. Jovovic, M. Draskovic, M. Delibasic, M. Jovovic. Journal of International Studies **10(1)** (2017)
11. E. Moallemi, S. Malekpour et al. One Earth **3(3)**, 300-313 (2020)
12. A. Kaminskyi et al., IOP Conf. Ser.: Earth Environ. Sci. **628**, 012031 (2021)
13. L. Zomchak, Yu. Drobotii. Katowice: PHUT. 20-26 (2020)
14. O. Pursky, A. Kiv, T. Dubovyk, I. Buchatska, H. Danylchuk. IOP Conference Series: Earth and Environmental Science **628(1)**, 012017 (2021).
15. V. Kadiyevskyy, N. Klymenko, Actual Problems of Economics **152(2)**, 313 (2014).
16. V. Babenko, I. Perevozova, M. Kravchenko, M. Krutko, D. Babenko, E3S Web Conf. **166**, 12001 (2020)
17. I. Semenenko, R. Halhash, K. Sieriebriak, Sustainable development of regions in Ukraine: before and after the beginning of the conflict. Equilibrium. Quarterly Journal of Economics and Economic Policy **14(2)**, 317 (2019)
18. P. Hryhoruk, N. Khrushch, S. Grygoruk, International Journal of Industrial Engineering and Production Research **31(4)**, 597 (2020)
19. N. Davydenko, Z. Skrypyk, O.V. Titenko, M. Zhovnirenko. Global Journal of Environmental Science and Management (**Special issue**) (2019)
20. Y. Starychenko, A. Skrypyk, V. Babenko, N. Klymenko, K. Tuzhyk. Studies of Applied Economics **38(4)** (2020)
21. M. Nehrey, I. Voronenko. European Journal of Molecular & Clinical Medicine **7(8)**, 5435 (2020)
22. V. Babenko, V. Sidorov, Y. Koniaieva, L. Kysliuk. Global Journal of Environmental Science and Management (**5**), 105 (2019)
23. D. Zherlitsyn, A. Skrypyk, N. Rogoza, S. Saiapin, T. Kudin, Estudios de Economia Aplicada **38(4)**, 3994 (2021)
24. Statistical Yearbook Environment of Ukraine 2018. http://www.ukrstat.gov.ua/druk/publicat/kat_u/2019/zb/11/Zb_dovk_2018.pdf. Accessed 21 Mar 2021