Application of Fuzzy Clustering to Shaping Regional Development Strategies in Ukraine

Kateryna Gorbatiuk
Department of Automated Systems and
Modeling in Economics
of Khmelnytskyi National University
Khmelnytskyi, Ukraine
kt_datsyuk@ukr.net

Oksana Proskurovych
Department of Automated Systems and
Modeling in Economics
of Khmelnytskyi National University
Khmelnytskyi, Ukraine
pov1508@gmail.com

Olha Mantalyuk
Department of Automated Systems and
Modeling in Economics
of Khmelnytskyi National University
Khmelnytskyi, Ukraine
olyasiko1@gmail.com

Oleksandr Valkov
Department of Automated Systems and
Modeling in Economics
of Khmelnytskyi National University
Khmelnytskyi, Ukraine
pulolalala@ukr.net

Abstract — **Formation of regional development strategies is impossible without taking into account the disparities in the development levels of regions which affect the entire national economy. Strengths and weaknesses of the functioning of each regional economic system, favourable and negative trends in the development of regions, disparities in the development levels, should be analysed and detected. This paper is devoted to applying of fuzzy clustering methods for the purpose of differentiation of development levels of the regions of Ukraine, in order to make recommendations on using of fuzzy clustering results in shaping regional development strategies. The fuzzy clustering methods were applied to the regional economic development indicators of the regions of Ukraine in 2017. There were considered different fuzzy clustering results and clustering validity techniques. The study presents the fuzzy classifications accompanied with the cluster validity process and substantial analysis of the economic indicators. The results of fuzzy clustering allow to consider in more detail the similarities in the regional economic development levels, assigned to the same clusters, and reveal the dissimilarities between the regions assigned to the different clusters. The membership coefficients can be used as values for regulation of influence on important economic components of the regional development strategies.**

Keywords — regional development disparities, regional development strategies, regional policy, clustering methods, fuzzy clustering, cluster validity.

I. INTRODUCTION

The effectiveness of the entire national economy heavily depends on the principles of regional policy, which form regional development strategies. Shaping the regional development strategies covers the issue of integration different development interests of the country, such as competitiveness of the regions and its strengthening on a country scale and on the international arena, as well as the prevention of excessive, socially and politically unacceptable differences between and inside the regions [1].

Shaping regional development strategies is seen as a better policy for delivering regional growth, which may represent a more efficient method of delivering regional development support, and lead to more effectiveness, in term of impact on job creation, entrepreneurship, enterprises activity and innovation [1]. Nowadays, regional policies are adapting to new challenges. They are going to focus on reducing disparities in income, infrastructure and on activities of a social nature. Regional policy directs the efforts at strengthening and using the endogenous potentials of all territorial and development mechanisms that support dissemination of development processes from the main growth centres to the whole regions [2]. This will enable the integration of the areas of whole regions into the process of development, competitiveness building and regional identity. Also, the regional policy further provides special activities directed at territories which require particular support to development processes, defined in the development strategies as problem areas. In this regard, the search for ways to minimize the negative effects of territorial inequalities of development, the development of effective regional policy is one of the central objectives in the countries of the world, especially in highly developed countries.

The application of scientific approaches to the problem of dividing Ukrainian regions into different groups accordingly to their economic development levels, based on different economic indicators, can elaborate economic policy measures for every region inside or outside groups. One of the approaches, which allows the gradual assessment of the membership of region in a cluster, described by a membership function valued in the real unit interval [0; 1], is fuzzy clustering technique. It provides an opportunity to identify soft boundaries for groups of regions, like in the case of most natural systems, and make it possible to more adequately describe regional development disparities.

This article presents a study on application of fuzzy clustering methods to problem of shaping regional development strategies in Ukraine though evaluating the development disparities in the regions of Ukraine. This approach is applied to the statistical data on indicators of economic activities in different regions of Ukraine in 2017, accompanied with the cluster validity process and substantial

analysis of the economic indicators, which we have taken as the base of the cluster research.

This article consists of six sections. The background of conducted research is presented in the first section. In the second section, there is a review of scientific literature on the research topic. The third section reveals the theoretical basis of the proposed techniques. The main results of the study are presented in the fourth and fifth parts of this paper. The final part contains conclusions based on the research results and recommendations on the use of fuzzy clustering methods for shaping regional development strategies.

II. LITERATURE REVIEW

Numerous international and European studies have shown that the regional policy models have produced disappointing results mainly due to their slow convergence [1]. Centrally managed redistribution of subsidies to less developed areas resulted in excessive dispersal of funds, thereby weakening the effect of the scale of public intervention, which ultimately could not produce measurable effects for development [2].

With the evolution of development challenges and economic, social and territorial conditions, the regional policy began to evolve from a model of short-term grants distributed by order from superior authority "for the most disadvantaged areas" to the model of long-term, decentralized development policies aimed at promoting all regions (irrespective of the degree of prosperity), from dispersed intervention to more selective investments [2]. Development strategies have begun to take into account a wide range of direct and indirect performance factors, focus on the endogenous territorial characteristics (instead of exogenous investments and transfers), putting emphasis on opportunities rather than barriers and on good governance, where an increasingly important role was played by regional and local authorities [2]. Also, such processes can be observed in Ukraine in recent years.

Examples of applying the regional development approach can be found in some countries (e.g. Sweden, Netherlands) [2]. Numerous scientific publications [1]-[5] suggest that development policies should support growth in all regions, and regions should invest in their own development by mobilizing local resources and funds in order to exploit their specific competitive advantages without excessive reliance on national transfers and grants [2]. Regional strategies should offer coherent methods of unifying regional resources and maximizing regional potential [1]. Local and regional actions will have to vary from one region to another to establish exceptional features for each region to support a decision-making process that can establish strategic priorities [1]. At the same time, the specific content of regional strategies is contingent with unequal economic development of the regions in different sectors [4], [5]. Consequently, the regional development strategies should that take into account differences in regional development, which can be assessed through specific scientific approaches to achieve maximum results in each region and for the whole country.

This research is an extension of the study [6] on the problem of clustering Ukrainian regions accordingly to their economic development levels. There were presented the new approach to assessing regional disparities in Ukraine using a fuzzy clustering technique. The statistical data on quantitative indicators of economic activities in different regions of Ukraine in 2017 were considered in [6], and it were used to obtain different partitions of the regions into fuzzy clusters.

The problems of estimating and classifying of regional development disparities, and some approaches to assessing the economic development levels, are discussed in [6]-[11]. Also, different clustering techniques, including fuzzy approaches, based on fuzzy sets theory of L. Zadeh [12], are presented in [13]-[18]. The theoretical basics on fuzzy clustering methods and their program software implementations, which we used in this research, are considered in [19]-[24]. In particular, several authors suggested some validation methods [25]-[27], which allow to reach the optimal partitions of the data.

III. RESEARCH METHODOLOGY

Fuzzy clustering is a soft clustering technique which allows an individual object to be partially classified into more than one cluster. In fuzzy clustering each object or data point belongs to all the clusters with varying membership values, which range between zero and one [17]. Consequently, the data objects closer to the centres of the clusters have higher degrees of membership than objects scattered in the borders of the clusters [19]. The advantage of this technique is that it does not assign every object into only one cluster, but in this case, the quantity of information for interpretation increases.

Generalization of the hard partition to the fuzzy case follows directly by allowing μ_{ik} to attain real values in [0, 1]. Conditions for a fuzzy partition matrix are given by (Ruspini, 1970) [17]:

$$\mu_{ik} \in [0,1], \quad 1 \leq i \leq c, 1 \leq k \leq N, \quad (1)$$

$$\sum_{i=1}^{c} \mu_{ik} = 1, \ 1 \leq k \leq N, \quad 0 < \sum_{i=1}^{c} \mu_{ik} < N, \ 1 \leq i \leq c. \quad (2)$$

where μ_{ik} – the membership function; c – the number of clusters; N – the number of data points.

Basic fuzzy clustering methods are: fuzzy clustering based on fuzzy relation, fuzzy clustering based on objective functions, and the fuzzy generalized K-nearest neighbour rule [6]. For most of fuzzy clustering algorithms, the number of clusters must be defined before applying the method. So, the quality of partitioning depends on the value of the number of clusters c. Therefore, the solving of the problem of assuming the clusters number parameter is one of the most important processes in clustering technique.

The problem of finding optimal number of clusters can be addressed by several cluster validity indices [25]-[27]. The number of cluster partitions, that can be obtained, are the same as the number of clusters centres c that is provided as an input to the clustering algorithms [25]. Wang et al. [27] have classified cluster validity algorithms into two types of categories, where the first category of algorithms uses only the membership matrix and the second category uses cluster centres and data points along with the membership matrix.

Also, we can use the silhouette analysis for the cluster validity purpose. It allows to study the separation distance

between the resulting clusters. In such case, we build the silhouette plot to display a measure, which has a range of [-1, 1], of how close each data point in one cluster is to points in the neighbouring clusters. So, it provides a way to assess clusters number parameter visually [19]. If silhouette coefficient is close to +1, then it indicates that the data point is far away from the neighbouring clusters. If a value of silhouette coefficient is close to 0, then it indicates that the data point is on or very close to the decision boundary between two neighbouring clusters. In the opposite case, a negative value of silhouette coefficient indicates that the data point might have been assigned to the wrong cluster.

Fuzzy clustering algorithms can be applied using the R software tools. We used the following R packages for our research: 1) cluster, ppclust and fclust for computing fuzzy clustering [20]-[22] and 2) factoextra for visualizing clusters [24]. Two functions fanny() (cluster package) and fcm() (pplcust package), were used to compute fuzzy partitions [18].

IV. CASE STUDY: DETECTING THE DEVELOPMENT DISPARITIES USING FUZZY CLUSTERING TECHNIQUES

Our research was conducted on the same data set as in [6]. The crisp clustering results and fuzzy clustering results are used for further study of development disparities of the regions of Ukraine to outline the ways of using this technique in shaping regional development strategies.

A. Data set description and feature selection

The statistic information about the economic activities in 2017 for regions, obtained from the State Statistics Service of Ukraine [28], was taken as the base of our study. The basic indicators of economic activities in the regions included the indicators of the extraction of aquatic bioresources, agriculture activities, retail trade, services and industrial activities [6]. They were divided into two groups, and the observations with the corresponding values were called the First data set and the Second data set. The description of these data sets along with their labels, used in the calculations output, are presented in the Tables I, II.

TABLE I. FIRST GROUP OF INDICATORS.

Indicator	Label
Extraction of aquatic bioresources by fishery water bodies	Bioresources
Value of agricultural products sold by agricultural enterprises	Products
The cost of sold agricultural products from plant growing	Plant
The cost of sold agricultural products from animal husbandry	Animal
Dynamics of sown areas of agricultural crops, all categories	Sown
Dynamics of sown areas of agricultural crops, agricultural enterprises	Enterprises
Dynamics of sown areas of agricultural crops, agricultural households	Households

TABLE II. SECOND GROUP OF INDICATORS.

Indicator	Label
Regional structure of turnover of retail trade	Retail
Completed construction works	Construction
Volume of industrial products sold	Products
Regional structure of retail trade turnover of retail enterprises	Enterprises
Used fuel, (10^3)	Fuel
Volume of services sold by enterprises in the service sector, (10^3)	Services

Firstly, we analysed the pairwise relationship between indicators in each data set by constructing the correlation plots using R software (Fig. 1, Fig. 2).
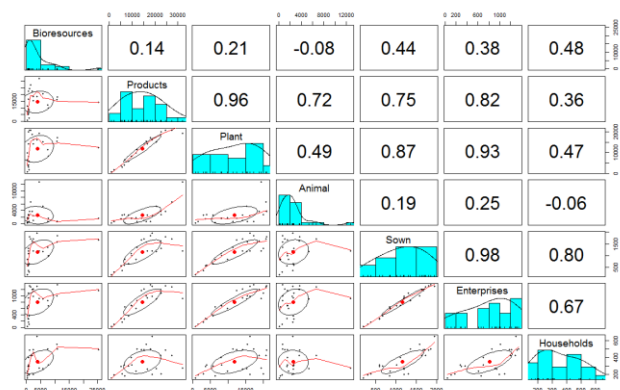

Fig. 1. Correlation plot for the First data set.

This plot shows that a lot of the indicators are weakly or not correlated (Bioresources and Products, Bioresources and Plant, Bioresources and Animal, Animal and Sown, Animal and Enterprises, Animal and Households). Some of them are moderately correlated (Bioresources and Sown, Bioresources and Enterprises, Bioresources and Households, Products and Animal, Products and Sown, Products and Households, Plant and Animal, Plant and Households, Enterprises and Households), and the others are strongly correlated (Products and Plant, Products and Enterprises, Plant and Sown, Plant and Enterprises, Sown and Enterprises, Sown and Households).
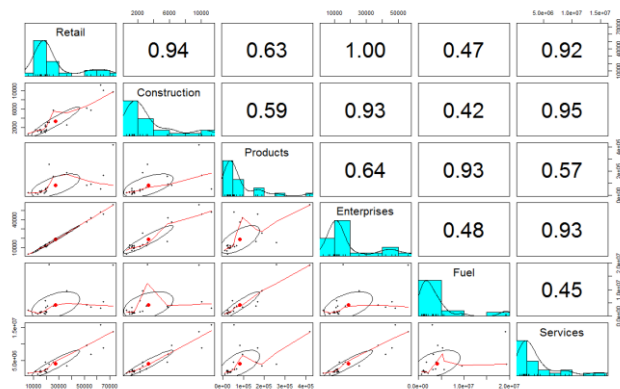

Fig. 2. Correlation plot for the Second data set.

As we can see, some of the indicators are strongly correlated (Retail and Construction, Retail and Enterprises, Retail and Services, Construction and Enterprises, Construction and Services, Products and Fuel, Enterprises and Services), the others are moderately correlated (Retail and Products, Retail and Fuel, Construction and Products, Construction and Fuel, Products and Enterprises, Products and Services, Enterprises and Fuel, Fuel and Services).

The correlation plots allowed us to conclude that some of the indicators should be removed from the First data set because they are not associated and cannot help to describe the common patterns in the data. The indicators, which should have been excluded, are Bioresources and Households. Therefore, we took for further work from the

First data set only the indicators: Products, Plant, Animal, Sown, Enterprises. Then, we selected all indicators from the second group for the fuzzy clustering of the regions. The values of these indicators, gathered in 2017, for all 24 administrative regions in Ukraine (Table III) were used for the fuzzy clustering.

TABLE III.   ADMINISTRATIVE REGIONS OF UKRAINE [6].

Id	Region	Id	Region
1	Vinnytsya	13	Mykolayiv
2	Volyn	14	Odesa
3	Dnipropetrovsk	15	Poltava
4	Donetsk	16	Rivne
5	Zhytomyr	17	Sumy
6	Zakarpattya	18	Ternopil
7	Zaporizhya	19	Kharkiv
8	Ivano-Frankivsk	20	Kherson
9	Kyiv	21	Khmelnytskiy
10	Kirovohrad	22	Cherkasy
11	Luhansk	23	Chernivtsi
12	Lviv	24	Chernihiv

Using these two sets of data with statistical values, we intended to get different partitions of the regions depending on the nature of corresponding economic indicators.

B. Fuzzy clustering

In [6], we applied the hierarchical clustering method by hclust() function [21], using a linkage method "single", to the both data sets. There we obtained the hard clusters for them, applied the fuzzy clustering methods, and obtained the fuzzy clusters, which were characterized by membership coefficients indicated the strength of belonging to the particular cluster for all regions [6].

In this research, we illustrate the fuzzy clusters obtained in [6] by several charts (Fig. 3-5) with the values of the membership coefficients obtained by the fcm() and fanny() functions. The number of clusters was four, and the values of membership coefficients vary from 0 to 1, that indicate the strength of belonging of a region to the particular cluster.
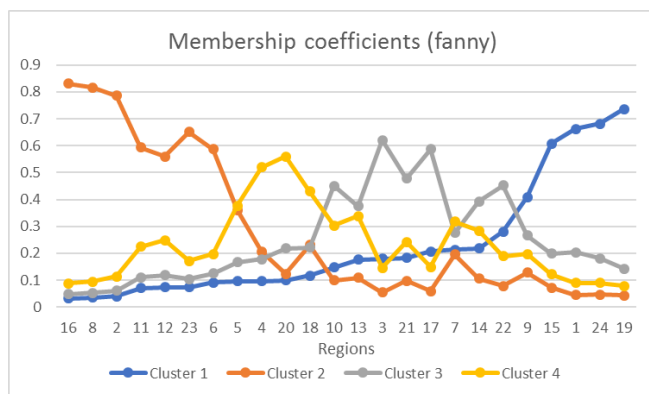

Fig. 3.   Membership coefficients by fanny() function for the First data set.

Different fuzzy clustering methods gives different results, but also there are common features in the fuzzy partitioning obtained by functions: fcm() and fanny(). Some of the regions are occur in the same cluster by both methods (for example, 16, 2, 8, 11, 12, 23, 6, 5 are in the Cluster 1 in Fig. 3 and in the Cluster 4 in Fig. 4).
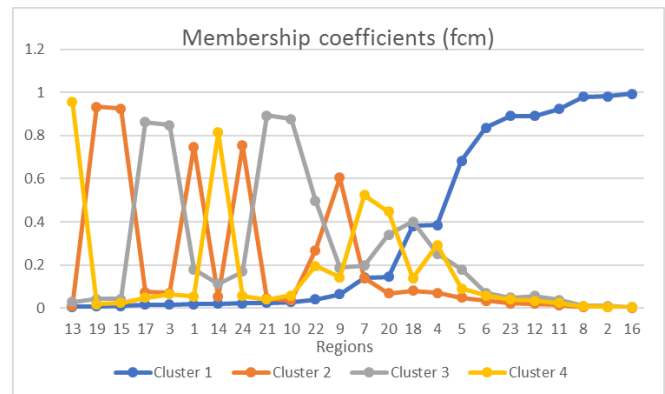

Fig. 4.   Membership coefficients by fcm() function for the First data set.

For the Second data set we obtained the chart of membership coefficients for four fuzzy clusters by fcm() function.
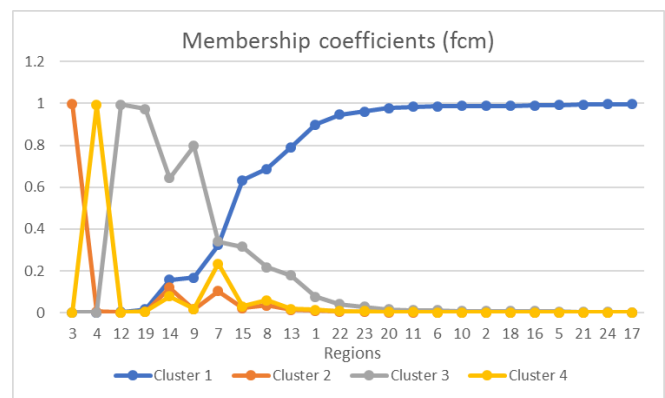

Fig. 5.   Membership coefficients by fcm() function for the Second data set.

On the plots above, the data points with the highest values of the membership coefficients can be joint into four different clusters to show which regions more likely are in the same cluster.

Finally, we created five clusters for the First data set because the most of the data points of the Second are very close to each other and cannot be split into different clusters. The results for the First data set and the number of clusters $c = 5$ are presented in the Fig. 6.
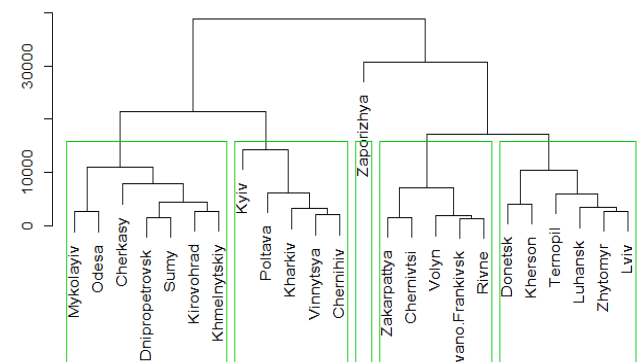

Fig. 6.   Five clusters for the First data set.

The plots on Fig. 3-5 show the overlapping clusters on the set of all regions. The data points with the membership function, which is close to one, belong to the same cluster. But some of the data points have the membership coefficients around 0.5. It means that the corresponding regions are not certainly belong to that cluster and could be assigned to another cluster as well. The closer the value of membership coefficient is to zero, the smaller the degree of belonging of a region to the particular cluster.

Five fuzzy clusters are also presented by different values of membership coefficients and overlapping ellipses (Fig. 7).
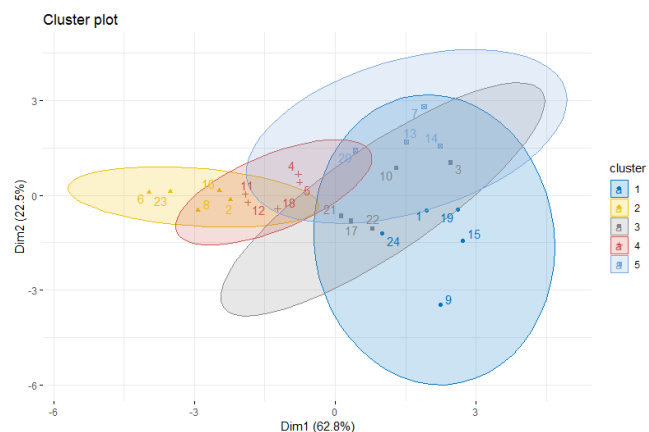

Fig. 7. Plot of the five fuzzy clusters for the First data set.

This is the scatterplot of the first two principal components which were derived from the data, which says that 85.3% (62.8%+22.5%) of the information about the multivariate data is captured by this plot.

The cluster validity process is the next task of our analysis. The main aim of it is finding the proper number of clusters in order to maximize the benefits out of clustering algorithms [25]-[27]. The optimal number of clusters is extremely significant because it strongly influences the quality of classifications. For this purpose, we applied the silhouette analysis to our fuzzy clustering results, and present the average silhouette coefficients for different number of clusters, both data sets and fuzzy clustering method fanny() in the Table IV.

TABLE IV.   SILHOUETTE COEFFICIENTS.

c	First data set	Second data set
2	0.57	0.58
3	0.49	0.56
4	0.44	0.51
5	0.35	0.48
6	0.36	0.26

As we can see, the partitions into two clusters have the largest values of the silhouette coefficient, and the optimal number of clusters for the First and Second data sets is two. But also, it could be three or four, no more.

Other validation techniques give us such indexes as: the fuzzy silhouette index (1), partition entropy (2), partition coefficient (3) and modified partition coefficient (4) [16] presented in the Table V.

TABLE V.   VALIDATION INDEXES.

c	First data set				Second data set			
	1	2	3	4	1	2	3	4
2	0.81	0.26	0.85	0.69	0.90	0.20	0.88	0.77
3	0.76	0.41	0.78	0.66	0.82	0.29	0.85	0.77
4	0.73	0.52	0.73	0.64	0.76	0.29	0.85	0.80
5	0.70	0.53	0.74	0.67	0.68	0.38	0.80	0.76
6	0.69	0.58	0.73	0.67	0.67	0.36	0.83	0.79

All these results also show that the optimal number of clusters is two for both data sets and for any clustering algorithm. Therefore, we can divide the Ukrainian regions into two groups with significant disparities in the economic development level. Each region in the clusters, obtained by different set of economic indicators, is described by the membership coefficient, which show the strength of belonging to this cluster, and could be used for differentiating the approaches to the development policy of the region.

C. Using fuzzy clustering results for shaping regional development strategies

In contemporary conditions, the regional policy must be focused on enabling a full utilisation of the development potentials of all regions of the country. Clustering results provide the processes of decision making in the regional policy with necessary information about the regional differentiations. In particular, the fuzzy clustering results give a measure for graduating the level of similarity of the common features of economic indicators in the groups of regions.

The membership coefficients in each fuzzy cluster could be standardised by the maximum value of all coefficients, and then, the greatest value of these standardised coefficients will be one:

$$\mu_{ik}^s = \frac{\mu_{ik}}{\max_k \mu_{ik}} \quad 1 \le i \le c, 1 \le k \le N . \quad (3)$$

The standardised coefficients μ_{ik}^s could be applied in the processes of implementation of the regional policy measures as the weight factor, which could balance the effect of the regional policy utilization in different regions within or outside the clusters. These coefficients also can be combined with composite index assessment technology, which is applied as the approach to determining the levels of economic development of the regions in Ukraine [11].

In order for the regional policy to be effective, it is necessary to define the development potentials for each region, with the simultaneous analysis of development factors, and on this basis, to precisely define the development opportunities for each region [2]. An important role in the generation of growth in the regions and their stable development is played by factors that can be considered as impacts on the regional economy.

The development potential in each region, in particular, can be assessed by applying the fuzzy clustering methodology, and then coming up with quantitative relationship between the membership coefficients of the regions in different fuzzy partitions and the values of impacts

on the regional economy. The membership coefficients can be used as values for regulation of influence on important economic components of the regional development strategies.

V. Results and Discussion

The results obtained in this study allows to develop a methodology for using the fuzzy clustering technique for shaping the development strategies through the assessing of the disparities in the economic development levels of the regions in Ukraine. The membership coefficients have a great potential to be used as parts of regulation factors in the regional economy because they provide the information on how different are the development levels within fuzzy clusters and between them.

The fuzzy clustering, along with the cluster validity techniques, and a substantial analysis of economic indicators, allow to justify the possibility of shaping the regional development strategies according to their different development levels, and use such partitions for regulation of the impact on important economic components.

VI. Conclusion

Shaping regional development strategies allows to deliver regional growth and lead to more effectiveness of entire national economy. Detecting the regional disparities helps to develop the effective economic policy for each region by taking action against the factors that slow down the economic growth, and maximizing the benefits of each region. Regional disparities in economic development level had been analysed in this study by different clustering techniques. Fuzzy clustering of the regions in Ukraine, based on two groups of economic indicators observed in 2017, allows to conclude that the information about the disparities across the regions can help to shape the right economic activities in each region, and support the decision-making processes in regional policies. The fuzzy clustering technique provides the instruments for creating the methodology of estimating the values of impact on regional economic activities in target sectors. In further research, it is necessary to pay attention to the quantitative assessment of the relationship between the values of membership coefficients and the values of impacts on the regional economy.

References

[1] J. Bachtler, and D. Yuill, *Policies and strategies for regional development: A shift in paradigm?* Glasgow, Scotland: University of Strathclyde, European Policies Research Centre, 2001.

[2] P. Zuber, *National Strategy of Regional Development 2010–2020: Regions, Cities, Rural Areas*. Warsaw: Ministry of Regional Development, 2010.

[3] I. Gill, "Improving Regional Development Policies," in: *OECD, Regional Outlook*, Paris: OECD, 2011.

[4] J. Stejskal, and P. Hajek, "Competitive advantage analysis: a novel method for industrial clusters identification," *Journal of Business Economics and Management*, vol. 13(2), 2012, pp. 344–365.

[5] K. Skokan, "The Role of Clusters in the Regional Policy of the Czech Republic," University Library of Munich, Germany, MPRA Paper 12353, 2007.

[6] K. Gorbatiuk, O. Mantalyuk, O. Proskurovych, and O. Valkov, "Analysis of regional development disparities in Ukraine with fuzzy clustering technique," in *Proceedings of 8th International Conference on Monitoring, Modeling & Management of Emergent Economy*, Odesa, 2019.

[7] Tb. Munandar, "Analysis of Regional Development Disparity with Clustering Technique Based Perspective," *International journal of advanced research in computer science*, vol. 6, Iss. 1, pp. 137–141, 2015.

[8] M. Lukovics, "Measuring Regional Disparities on Competitiveness Basis," in: Z. Bajmocy, I. Lengyel, eds., *Regional Competitiveness, Innovation and Environment*, JATEPress, Szeged, 2009, pp. 39–53.

[9] V. Hryhorkiv, A. Verstiak, O. Verstiak, and M. Hryhorkiv, "Regional Economic Growth Disparities in Ukraine: Input-Output Analysis Approach," *Scientific Annals of Economics and Business*, vol. 64(4), pp. 447–457, 2017.

[10] S. Maksymenko, "Ukraine's Regional Economic Growth and Analysis of Regional Disparities," Working Paper 6053, Department of Economics, University of Pittsburgh, 2016.

[11] P. Hryhoruk, N. Khrushch, and S. Grygoruk, "An approach to design a composite index of economic development and identifying the bounds of its levels," in *Proceedings 9th International Conference on Advanced Computer Information Technologies ASIT'2019*, Ceske Budejovice, 2019, pp. 48-51

[12] L.A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, pp. 338–358, 1965.

[13] A. Gosain, and S. Dahiya, "Performance Analysis of Various Fuzzy Clustering Algorithms: A Review," *Procedia Computer Science*, vol. 79, 2016, pp. 100–111.

[14] D.J. Miller, C. Nelson, M.B. Cannon, and K.P. Cannon, "Comparison of Fuzzy Clustering Methods and Their Applications to Geophysics Data," *Applied Computational Intelligence and Soft Computing*, 2009, [Online]. Available: https://www.hindawi.com/journals/acisc/2009/876361/. Accessed June 19, 2019.

[15] V.K. Malhotra, H. Kaur, and M.A. Alam, "An Analysis of Fuzzy Clustering Methods," *International Journal of Computer Applications*, vol. 94, no. 19, pp. 9–12, 2014.

[16] M.-S. Yang, "A survey of fuzzy clustering," *Mathematical and Computer Modelling*, vol. 18, no.11, pp. 1–16, 1993.

[17] E.H. Ruspini, "Numerical methods for fuzzy clustering," *Information Sciences*, vol. 2, no. 3, pp. 319–350, 1970.

[18] V. Babenko, I. Perevozova, O. Mandych, T. Kvyatko, O. Maliy, and I. Mykolenko, "World informatization in conditions of international globalization: factors of influence," *Global. J. Environ. Sci. Manage.*, vol. 5(SI), pp. 172-179, 2019.

[19] RPubs. Unsupervised Learning – Clustering Fuzzy C Means. [Online]. Available: http://rpubs.com/ rahulSaha/Fuzzy-CMeans Clustering. Accessed on: June 19, 2019.

[20] M.B. Ferraro, and P. Giordani, A toolbox for fuzzy clustering using the R programming language. *Workshop on Clustering methods and their applications.* [Online]. Available: http://pro1.unibz.it/projects/Clustering_Methods_2014/Ferraro.pdf. Accessed on: May 18, 2019.

[21] M. Maechler, P. Rousseeuw, A. Struyf, and M. Hubert, Package 'cluster'. CRAN. [Online]. Available: https://cran.r-project.org/web/packages/cluster/cluster.pdf (2018). Accessed on: June 19, 2019.

[22] P. Giordani, M.B. Ferraro, and A. Serafini, Package 'fclust'. CRAN. [Online]. Available: http://cran.r-project.org/web/packages/fclust/fclust.pdf. Accessed on: June 19, 2019.

[23] Z. Cebeci, F. Yildiz, A.T. Kavlak, C. Cebeci, and H. Onder, Package 'ppclust'. CRAN. [Online]. Available: https://cran.r-project.org/web/packages/ppclust/ppclust. Accessed on: June 23, 2019.

[24] A. Kassambara, and F. Mundt, Package 'factoextra'. CRAN. [Online]. Available: https://cran.r-project.org/web/packages/factoextra/factoextra.pdf. Accessed on: May 28, 2019.

[25] S. Joopudi, S.S. Rathi, S. Narasimhana, and R. Rengaswamyb, "A New Cluster Validity Index for Fuzzy Clustering," *IFAC Proceedings Volumes*, vol. 46(32), 2013, pp. 325–330.

[26] J.C. Bezdek, "Cluster validity with fuzzy sets," *Journal of Cybernetics*, vol. 3, no. 3, 58–73, 1974.

[27] W. Wang, and Y. Zhang, "On fuzzy cluster validity indices," *Fuzzy Sets Syst*, vol. 158, no. 19, pp. 2095–2117, 2007.

[28] State Statistics Service of Ukraine. [Online]. Available: http://www.ukrstat.gov.ua. Accessed on: February 20, 2019.