

MAŁGORZATA MARKOWSKA, BARTŁOMIEJ JEFMAŃSKI

FUZZY CLASSIFICATION OF EUROPEAN REGIONS IN THE EVALUATION OF SMART GROWTH¹

1. INTRODUCTION

It is possible to accomplish smart growth only if an overall innovation potential of EU regions becomes mobilized, since innovation is of great importance for all regions – for the well developed ones, to maintain the role of leaders, and also for these lagging behind to be able to catch up with the best ones [17]. Regions play the crucial role in smart growth accomplishment and in this case smart specialization profile is of decisive importance since it is the regions which take up the role of main institutional partners for universities, other research and educational institutions, as well as small and medium enterprises constituting the key component of innovation processes responsible for overall development [23].

Complex phenomena, such as: development, innovation or smart specialization in regions represent processes difficult for quantification and, in consequence, for measurement. Therefore for the purposes of their identification and characteristics it seems founded to apply tools used in contemporary econometrics, including multidimensional classification methods, the results of which allow for e.g. performing comparative and benchmarking analyses between spatial units, i.e. regions. Such approach has also been applied in the hereby study, which focuses on the evaluation of smart growth level in European regional space. The study applies European space division into fuzzy classes using the output of fuzzy sets theory and therefore allows the option of regions partial membership in particular classes. Substantiality of such approach was characterized, e.g. in the study by Y. Leung [14] who emphasized that numerous spatial analyses were carried out following the assumption that a region represents a strictly defined theoretical construct, characterized by clearly outlined spatial boundaries. This approach, on the other hand, implies the division of space into mutually exclusive regions, in line with Boole's logics. The reality, however, shows that in case of the majority of special phenomena it is not possible to define explicit boundaries between regions, which questions the reasons behind applying, let's call it, dichotomous regionalization.

¹ The study was prepared within the framework of research grant provided by The National Centre of Science no.2011/01/B/HS4/04743 entitled: *Classification of the European regional space in the perspective of smart development concept - dynamic approach*.

Therefore, also in the hereby study, the authors decided to take advantage of fuzzy sets theory output and, in particular, of fuzzy algorithms classification methods. The obtained results indicated both correctness and usefulness of the used approach and confirmed the assumption adopted by the authors that the boundaries between distinguished classes are not clear and in case of some regions it is difficult to indicate their unambiguous membership to the defined classes.

2. SMART SPECIALIZATION AS THE PILLAR OF SMART GROWTH

2.1 SMART GROWTH IN EU STRATEGIC DOCUMENTS

In the perspective of minor effects resulting from Lisbon Strategy implementation and in view of global economic crisis the united Europe needs a strategy which can open opportunities for coming out of the crisis not only stronger, but which may also result in EU economy becoming both smart and sustainable, supporting social inclusion, obtaining high employment rate and capacity indicators, as well as more extensive social cohesion. Therefore it seems crucial to release European innovation potential by improving the effects of education process as well as the quality and results of educational institutions functioning and also by using economic and social opportunities created by a digital society. Impulses should refer to regional, national and EU level.

In times of global crisis and in view of growing competition at an international arena, e.g. on the part of developed and emerging countries and also the need to reform global financial system, as well as challenges determined by climate changes and natural resources, EU needs new strategy based on coordinated economic policy, the strategy which ensures both development and employment rate growth [6].

In view of such challenges the European Commission, on 3rd March 2011, presented the communication *Europe 2020 – Strategy for smart growth and sustainable development enhancing social inclusion* [8]. The Commission proposal regarding new strategy initiation was accepted on 26 March 2010 at The European Council meeting. *Europe 2020* strategy, as the successor of Lisbon Strategy, represents the vision of social market economy for Europe of the 21st century and refers to three related priorities [8]:

- smart growth: knowledge-based economy and innovation development;
- sustainable growth: support for effective economy which uses resources, more environmentally friendly and more competitive one;
- inclusive growth: support for economy representing high employment rate, ensuring social and territorial cohesion.

EU strategic documents define smart growth in terms of obtaining better results in:

- education, by encouraging towards learning, studying and upgrading qualifications,
- research and innovation, by creating new products and services influencing economic growth and increasing employment rate which facilitate solving social problems,

- digital society, i.e. information and communication technologies implementation.
- Guideline initiatives stimulating smart growth in EU are as follows:
- The European digital agenda [1, 5], i.e. establishing uniform digital market based on very fast Internet connections and interoperation applications;
 - Innovation Union [11], implementing R&D and innovation projects in solving crucial economic problems; strengthening both the process and the role of innovation;
 - Mobile youth, i.e. “Youth on the move” project, aims at improving educational results and increasing the attractiveness of European college education at an international arena.

Other initiatives (flagships in Europe 2020 strategy) are as follows [8]:

- Europe using its resources in an effective way,
- Industrial policy in globalization era,
- Programme for new skills and employment,
- European programme for fighting poverty.

Each of the listed above initiatives was also accompanied by measurable and possible to accomplish targets.

2.2 SMART SPECIALIZATION AS THE COMPONENT STIMULATING SMART GROWTH

Europe 2020 strategy represents an integrated approach where, apart from sustainable growth related priorities, emphasis is also placed on inclusive growth and smart growth. Smart growth, in the perspective of strategic documents [8], knowledge-based economy and innovation, means an extended role of knowledge and innovation as the forces stimulating future growth, which requires improvements in the quality of education, higher efficiency of research projects, supporting innovation and knowledge transfer in EU, better implementation of information and communication technologies, as well as ensuring that innovative ideas find their reflection in new products and services which stimulate better growth rate dynamics, facilitate opening new jobs and enhance solving crucial social problems in Europe and worldwide.

The success of Europe 2020 Strategy also depends on such elements as entrepreneurship, financial resources and paying attention to both, users’ needs, and opportunities offered by the market.

Smart specialization constitutes an important component of smart growth, which covers enterprises, research centres and higher education institutions cooperating in order to define the most promising areas of specialization in a given region, but also weaknesses making innovation implementation difficult. This approach takes into consideration differences in business opportunities specific for particular regions with reference to innovation, since while the leading regions can invest in upgrading general technologies or innovations in services, in case of the other regions better results may be obtained by investing in innovations in a specific sector, or a few related sectors.

Smart specialization accomplishment turns out possible, if regional diversity is taken advantage of, by stimulating cooperation and extending it beyond (regional and

national) boundaries, by regional opening towards new opportunities, by avoiding fragmentation and as the result offering undisturbed knowledge transfer in EU.

Strategic intelligence which allows for defining activities of high added value and therefore offers maximum opportunities for upgrading regional competitiveness, turns out necessary for smart growth accomplishment. Significant influence of investments in R&D and innovation requires obtaining critical mass which should be accompanied by funds aimed at upgrading skills, as well as an overall level of educational institutions and infrastructure extension. State governments (national level) and local governments (regional level) should develop strategies for “smart specialization” in order to obtain optimum effects of the carried out regional policy and other EU policies.

3. FUZZY CLASSIFICATION

Fuzzy set \tilde{A} in $X = \{x\}$ space marked as $\tilde{A} \subseteq X$ is defined by the set of pairs $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\} \forall x \in X$, where $\mu_{\tilde{A}} : X \rightarrow [0, 1]$ means the membership to \tilde{A} fuzzy set function, which assigns to each $x \in X$ element its membership level to \tilde{A} fuzzy set, [26 pp. 11-12]. Therefore, a fuzzy set represents the generalization of a classical set, where element belongs ($\mu_{\tilde{A}}(x) = 1$) or does not belong ($\mu_{\tilde{A}}(x) = 0$) to A set [24, p. 100].

In classical classification methods an object’s membership in a class is expressed by a binary variable. In other words, either an object belongs to a given class or it does not. The application of fuzzy sets concept in the problem of classification allows for the option of an object’s membership in more than just one class. It is possible due to binary variable substitution by a continuous variable presenting values in $[0; 1]$ interval. This procedure allows for describing the situation more precisely, i.e. where the boundaries between classes are “unclear” and assigning an object to a unique class becomes more difficult [12, p. 160]. It also reflects the actual reality to a higher extent and may serve as the protection for a research worker preventing him/her from losing certain information, as opposed to an approach where objects are assigned to one class only [13, p. 146-156].

The profile of fuzzy classification methods has already been discussed in studies, e.g. by: F. Höppner [10], K. Jajuga [12] and F. Wysocki [24]. One of the more frequently applied methods is the fuzzy c -means method suggested by A. Dunn [7] and later generalized by J. C. Bezdek [2] and F. Höppner [10, p. 37]. An interesting and comprehensive review of k -means method development through its diversified modifications and generalizations (including fuzzy classification) was presented by Bock [3]. An extensive description of this method covering its advantages and disadvantages, was also included in the study by Steinley [20]. Its application does not require making assumptions regarding the nature of empirical material under analysis. It is an iteration method the idea of which is very close to classical k -means method. The method focuses on finding such gravity centres of classes which minimize the function below

[15, p. 303]:

$$J_m = \sum_{i=1}^c \sum_{j=1}^n \mu_{ij}^m d_{ij}^2, \quad (1)$$

where: μ_{ij} – level of j -th object membership in i -th fuzzy class,

d_{ij} – Euclidean distance between i -th fuzzy class gravity centre and j -th object,

m – fuzzification parameter, where $m > 1$.

The algorithm of fuzzy c -means method consists of the following steps [4, pp. 230-238]; [15, pp. 302-303]:

Step 1. Random membership matrix initiation $\mathbf{U} = [\mu_{ij}]$ where: $\sum_{i=1}^c \mu_{ij}, \forall j = 1, 2, \dots, n$.

Step 2. Calculating gravity centres of classes following the below formula:

$$c_i = \frac{\sum_{j=1}^n \mu_{ij}^m x_j}{\sum_{j=1}^n \mu_{ij}^m}, \quad (2)$$

where:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{m-1}}}. \quad (3)$$

Step 3. Calculating new \mathbf{U}_{new} membership matrix. If $\|\mathbf{U}_{new} - \mathbf{U}\| > \varepsilon$, where $\|\mathbf{U}_{new} - \mathbf{U}\|$ represents Euclidean distance while ε refers to the accepted convergence threshold, then $\mathbf{U} = \mathbf{U}_{new}$ should be accepted and we should proceed to step 2. Procedure is finalized in the situation when $\|\mathbf{U}_{new} - \mathbf{U}\| < \varepsilon$ or the set number of k iterations is obtained.

Before initiating calculations the applied distance measure, number of classes, initial membership of objects in classes and the value of m fuzzy parameter, have to be defined.

M parameter specifies the degree of classification results fuzzification. Parameter value should be $m > 1$ where values close to unity will bring about the results similar to these received by means of classical methods. An increase in m parameter value results in the fact that the degree of objects membership in particular classes will take values close to inverse number of classes, i.e. $\frac{1}{c}$ [3, pp. 228-229]; [13, p. 146]. Professional literature does not offer theoretical background for the choice of optimum m parameter value, therefore its choice is frequently made based on experience from previously conducted empirical studies. F. Wysocki's research results suggest that parameter value should be included in [1,3;1,5] interval [24, p. 101].

4. CHARACTERISTICS OF EU REGIONS IN TERMS OF SMART SPECIALIZATION IDENTIFIERS

The implementation of visions described in the strategy requires specifying targets, therefore the targets set for ensuring smart growth are as follows:

1. higher level of outlays on investments (public and private) up to 3% of EU GDP and providing better conditions for R&D and innovation,
2. higher share of the employed (women and men) aged 20-64 up to 75% in 2020 as the result of introducing more people to the job market (mainly women, young and senior citizens, unqualified individuals and legal immigrants),
3. higher level of education, especially by:
 - reducing the share of young people finishing their education prematurely to less than 10%,
 - increasing the share of university graduates aged 30-34 up to at least 40%.

Measurable targets were also defined for other strategy elements, i.e. sustainable growth and inclusive growth. Even though so much has been written about smart specialization of regions, still the smart specialization identifying parameters have not been defined.

Specialization results in different effects depending on technological level it refers to, while the range of its occurrence may be indicated, as below [9]:

- specialization in the domain of scientific knowledge,
- specialization in technology and innovation,
- specialization in production processes,
- specialization related to clusters,
- horizontal vs. vertical specialization.

In professional literature specialization measurement originates from trade theory. Diversified specialization indicators were prepared in order to allow for capturing country specialization, but also different indicators were elaborated and used as technology specialization indicators and obviously, after certain corrections were applied.

The growing service orientation is one of the first specialization symptoms, therefore it seems that defining the importance of service sector at the background of other sectors in economy will be a helpful indicator in defining smart specialization. The service sector may be defined by:

X_1 – share of workforce employed in farming in an overall number of workforce in the region,

X_2 – share of workforce employed in industry in an overall number of workforce in the region,

X_3 – share of workforce employed in services in an overall number of workforce in the region.

Another manifestation of smart specialization is the growing importance of innovative sectors which may be assessed by means of the following characteristics:

X_4 – workforce employed in knowledge-intensive services as the share of overall workforce in services,

X_5 – workforce employed in high and mid-tech industry (as % of all workforce employed in industry).

The list of due characteristics could be much longer, however, in many cases regional level data are not available for properties potentially helpful in smart specialization identification.

The identification of smart specialization was prepared based on the set of EU NUTS 2 regions, the total number of which amounts to 271 [18], however, due to data gaps referring to selected characteristics about French overseas regions (Guadeloupe, Martinique, Guyane, Réunion) and two Spanish ones (Ciudad Autónoma de Ceuta, Ciudad Autónoma de Melilla) further analysis covers 265 out of 271 EU regions. The data regarding regions were collected from Eurostat data base and referred to 2008.

Measures selected for the evaluation of smart specialization, i.e. the five mentioned above characteristics are most diversified (in assessing the variability and also maximum and minimum ratio) regarding the share of workforce employed in farming in the total number of workforce in the region (variability coefficient equals 115% and the max/min ratio amounts to as much as 303), while the least diversified in case of two properties:

- share of workers employed in services in the total number of workers in the region (variability coefficient equals 15,9% and max/min ratio is 3),
- workforce employed in knowledge-intensive services as the share of total workforce in services (variability coefficient – 15,9%, max/min ratio – 2,9).

Extreme values for particular variables were as follows:

- share of workforce employed in farming in the total number of workforce in the region (from 0,16% up to 48,5%) and among regions where the property value was below 1%, 12 regions were listed (Utrecht (NL), Hovedstaden (DK), Wien (AT), Outer London (UK), Berlin (DE), Région de Bruxelles (BE), Bremen (DE), Hamburg (DE), Stockholm (SE), Inner London (UK), Île de France (FR), Praha (CZ), while among regions characterized by the share of workforce employed in farming above 25% 11 regions were included (Nord-Est (RO), Sud-Vest Oltenia (RO), Lubelskie (PL), Sud – Muntenia (RO), Świętokrzyskie (PL), Podlaskie (PL), Peloponnisos (GR), Sud-Est (RO), Nord-Vest (RO), Podkarpackie (PL), Anatoliki Makedonia, Thraki (GR)),
- share of workforce employed in industry sector in the total number of workforce in the region (9,4% – 46,8%), out of which in thirteen regions it was less than 16% (four Dutch regions (Flevoland, Zuid-Holland, Noord-Holland and Utrecht), two British ones (Outer London and Inner London) and Greek (Voreio Aigaio, Ionia Nisia), and also Région de Bruxelles (BE), Hovedstaden (DK), Île de France (FR), Stockholm (SE) and Luxembourg), while in fifteen regions above 40% (6 Czech regions: Severovýchod, Střední Morava, Jihozápad, Jihovýchod, Střední Čechy, Moravskoslezsko, two from Romania (Centru, Vest), two from Slovenia (Západné Slovensko, Stredné Slovensko) and two from Hungary (Közép-Dunántúl, Nyugat-

- Dunántúl) and also one Italian (Marche), German (Stuttgart) and Spanish region (La Rioja)),
- workforce employed in services in the total number of workforce in the region (30,4% – 90,2%), out of which above 80% in 19 regions (4 British (Inner London, Outer London, Berkshire, Bucks and Oxfordshire, Surrey, East and West Sussex) and four Dutch (Flevoland, Utrecht, Noord-Holland, Zuid-Holland), two Belgian (Région de Bruxelles, Prov. Vlaams Brabant), German (Berlin, Hamburg) and French (Île de France, Languedoc-Roussillon) and also Wien, Praha, Danish Hovedstaden, as well as Stockholm and Luxembourg), while below 50% in 18 regions (eight Polish: Małopolskie, Lubelskie, Podkarpackie, Świętokrzyskie, Podlaskie, Wielkopolskie, Opolskie and Kujawsko-Pomorskie, seven Romanian: Nord-Vest, Centru, Nord-Est, Sud-Est, Sud – Muntenia, Sud-Vest Oltenia, Vest and also Severovýchod (CZ), Centro (PT) and Vzhodna Slovenija (SI)),
 - workforce employed in knowledge-intensive services as the share of total workforce employed in services (25,2% – 73,5%), where above 60% was registered in 14 regions (8 Swedish: Stockholm, Östra Mellansverige, Småland med öarna, Sydsverige, Västsverige, Norra Mellansverige, Mellersta Norrland, Övre Norrland, three British: Inner London, Berkshire, Bucks and Oxfordshire, Surrey, East and West Sussex, two Finnish (Pohjois-Suomi, Åland) and Danish Hovedstaden), while below 35% in 13 regions (7 Greek: Anatoliki Makedonia, Thraki, Ionia Nisia, Sterea Ellada, Peloponnisos, Voreio Aigaio, Notio Aigaio, Kriti, two Portuguese (Algarve, Região Autónoma dos Açores) and Romanian (Centru, Sud-Est) and also Bulgarian (Yuzhen tsentralen) and Spanish (Canarias)),
 - workforce employed in high and mid-tech industry sectors (as % of workforce in industry) – from 3,9% to 59%, including 10 regions characterized by the property value below 6% (7 Greek regions: Anatoliki Makedonia, Thraki, Dytiki Makedonia, Ipeiros, Dytiki Ellada, Voreio Aigaio, Notio Aigaio, Kriti, two Spanish (Extremadura, Canarias) and Cyprus), similarly to the situation of regions characterized by the property value above 40% where eight German regions were listed: Stuttgart, Karlsruhe, Freiburg, Tübingen, Oberbayern, Bremen, Braunschweig, Rheinhessen-Pfalz and two French ones (Alsace and Franche-Comté).

The highest (negative) correlation was registered between X_1 and X_3 properties (-0,75) and X_2 and X_3 (-0,72) and also X_1 and X_4 (-0,47), while the next (positive) correlation between the following properties: X_3 and X_4 (0,53) as well as X_4 and X_5 (0,44) which means that an increase (decrease) of workforce share in farming is accompanied by a drop (growth) of workforce share employed in services and knowledge-intensive services, while an increase (decrease) of workforce share in industry goes along with the drop (growth) of workforce share in services. The same orientation is, however, observed in case of:

- changes in the share of workforce employed in services and knowledge-intensive services, as well as

- changes in the share of workforce employed in knowledge-intensive services in high and mid-tech industry sectors.

5. CLASSES OF REGIONS OBTAINED BY APPLYING FUZZY *C*-MEANS METHOD

The application of fuzzy *c*-means method in order to distinguish classes of regions, just like in case of classical *k*-means method variant, requires the number of such classes to be specified. If the non-statistical information about the number of classes is missing two solutions are suggested [24, p. 136]:

- to accept the number of classes specified using disjoint classification methods for the same data matrix,
- to perform fuzzy classification for a different number of classes and select the one for which fuzzy classification quality index reaches an extreme level.

In the hereby study the first procedure was accepted. As Sokołowski [19] emphasizes no classification method exists the advantage of which over other methods would be commonly acceptable. Therefore it is suggested to apply different methods in solving the classification problem and compare the obtained results. In view of the above, in the hereby analysis three, well recognized in professional literature, classification methods were applied in order to divide regions into classes: Ward method, *k*-medoid method and *k*-means method in relation to *Gap* classification quality assessment index. In order to standardize values of all variables, the regarding their variability, the standardization formula was applied. In this way variability, as the basis for objects diversification, was eliminated. Detailed characteristics and properties of variable normalization formulas (including standardization) may be found, among others, in studies by Pawełek [16] as well as Walesiak and Gatnar [22]. The similarity of objects was assessed by means of Euclidean distance (in case of Ward method square of Euclidean distance was applied). *Gap* index values and *diffu* statistics for a different number of classes, specified by means of the three classification methods, are illustrated in Table 1.

Graphical presentation of *Gap* index for a different number of classes is shown in picture 1, while the results of classification obtained by means of Ward method are presented in a dendrogram form in picture 2.

The results of classification obtained using Ward method and *k*-medoid method suggest the division of regions into four classes. In case of *k*-means method *Gap* index values suggest the division into five classes. The assessment of classification stability level in case of four and five classes variant was conducted by means of replication analysis [22, p. 419-420]. The analysis results are presented in Table 2.

The highest stability is observed in the variant of regions classification into four classes using *k*-means method. Classification stability in case of the suggested by *Gap* index division into five classes, and applying *k*-means method, proved lower. High stability is also characteristic for the division into four classes by means of the other

Table 1.

Index values of *Gap* and *diffu* statistics classification quality assessment depending on the number of classes

Class no.	Classification methods					
	Ward		<i>k</i> -medoid		<i>c</i> -means	
	<i>Gap</i>	<i>diffu</i>	<i>Gap</i>	<i>diffu</i>	<i>Gap</i>	<i>diffu</i>
2	1,01530	-0,04552	1,04846	-0,07183	1,0557	-0,05586
3	1,09024	-0,07141	1,15610	-0,01635	1,13291	-0,03074
4	1,18634	0,01484	1,21552	0,06360	1,21083	-0,00120
5	1,20840	0,04596	1,18817	0,05279	1,24119	0,06969
6	1,21571	0,05553	1,19104	0,04110	1,21806	0,03111
7	1,17350	-0,01324	1,16899	0,05565	1,21732	0,04697
8	1,20252	0,03294	1,12882	-0,06117	1,18542	-0,02165
9	1,20446	0,00085	1,22515	0,00730	1,20594	-0,01057
10	1,20486	0,01770	1,21747	0,01469	1,24870	0,06138

Source: Author's compilation using clusterSim package in **R** programme.

Table 2.

Adjusted Rand measure values

Class no.	Adjusted Rand measure		
	Ward	<i>k</i> -medoid	<i>c</i> -means
4	0,648875	0,736938	0,7411473
5	0,6402254	0,6096901	0,4975753

Source: Author's compilation using clusterSim package in **R** programme.

two methods. Owing to the above results, related to quality and stability of the obtained classification, for the purposes of fuzzy classification, the division of 265 EU NUTS 2 level regions into four classes was performed.

Fuzzy classification applying fuzzy *c*-means method requires, apart from *a priori* specification of classes number, also the initial classification of objects. Possible approaches in this matter are presented, inter alia, in F. Wysocki's study [24, p. 102]. The hereby analysis applies one of them which consist in random assignment of objects into four classes.

The similarity of objects was assessed by means of Euclidean distance. Fuzzy parameter value was set at the level of $m = 1,5$. Gravity centres for the distinguished classes are presented in Table 3.

The first class is characterized by the highest average share of workforce employed in farming sector in the total number of workforce in the region (24,3%) and the second

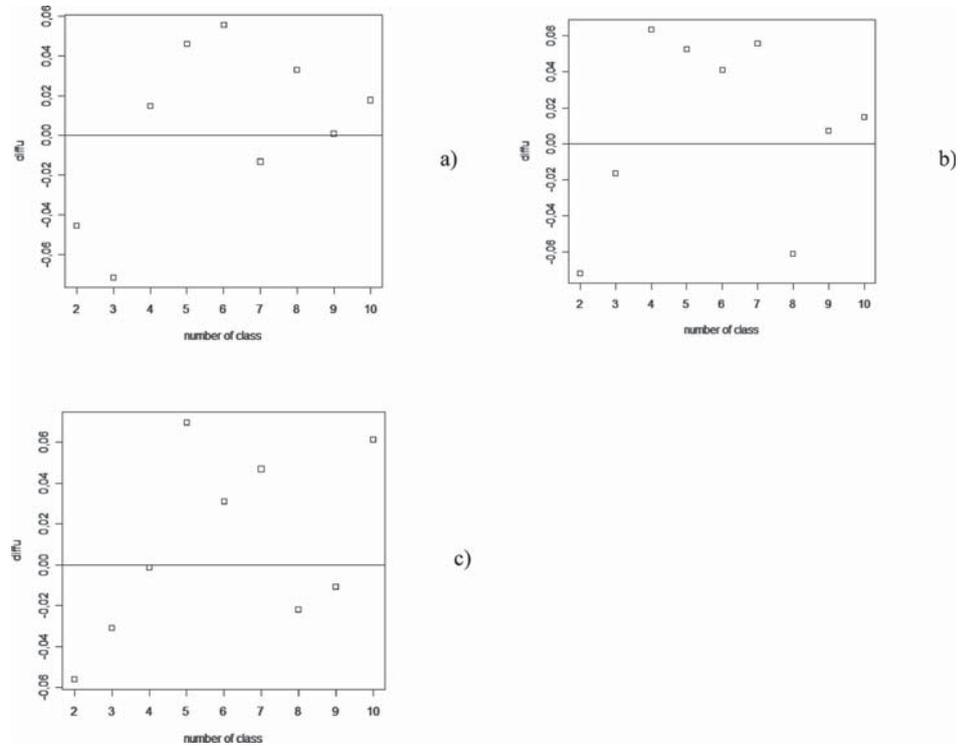


Figure 1. *Gap* index values graphs respectively for the following methods: Ward, *k*-medoid and *c*-means

Source: Author's compilation using clusterSim package in R programme.

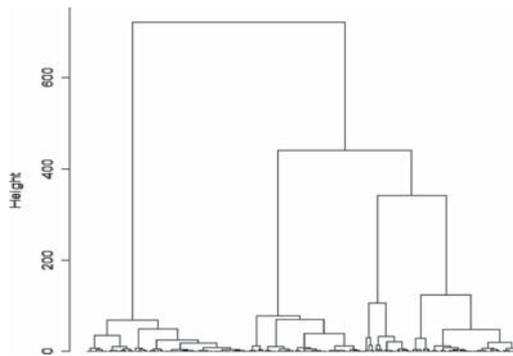


Figure 2. Dendrogram prepared using Ward method

Source: Author's compilation using clusterSim package in R programme.

largest average share of workforce employed in industry sector (28,8%), and the lowest average share of:

- workforce employed in services in the total number of workforce in the region (46,98%),

Table 3.

Arithmetic means for four classes obtained using fuzzy *c*-means method and for classes obtained at threshold membership value of 0,8

Class no.	variable					variable				
	X_1	X_2	X_3	X_4	X_5	X_1	X_2	X_3	X_4	X_5
1	24,26	28,75	46,98	40	13,5	26,02	26,83	45,78	40,37	12,64
2	6,83	26,25	66,92	41,7	15,3	6,71	25,72	67,83	41,57	14,18
3	4,62	35,72	59,66	47,4	31,9	4,05	36,48	59,14	46,74	31,27
4	2,82	21,79	75,38	54,4	25,3	2,12	21,55	75,97	54,73	25,24

Source: Author's compilation using **R** programme.

- workforce employed in knowledge-intensive services in the total number of workforce employed in services in the region (40%),
- workforce employed in high and mid-tech industry sector in the total number of workforce in industry (13,5%).

The first class covers the smallest number of regions, i.e. only 10,2%, i.e. 27 European space regions (9, i.e. 4,3% of EU 15 regions and 18, i.e. 32% of EU 12 regions) and agglomerates NUTS 2 level units representing 5 countries: 10 out of 16 Polish ones, 7 out of 8 Romanian, 7 out of 13 Greek, 2 out of 7 Portuguese and 1 out of 6 Bulgarian regions.

The second class, covering 70 regions, i.e. every fourth EU region, is characterized by the second largest average share of workforce employed in farming (6,8%) and services (67%) and the third highest average level of X_2 , X_4 and X_5 properties. The second class group of regions includes:

- 57 EU 15 regions (27,3%), including: 13 from 17 Spanish (76,5%) regions, 12 from 21 Italian (57,1%), 6 from 9 Austrian (66,7%), 6 from 22 French (27, 3%), 6 from 13 Greek (46,2%), 5 from 7 Portuguese (71,4%), 4 from 39 German (10,3%), 2 from 31 British (5,4%) and one from Belgium, Ireland and The Netherlands,
- 13 EU 12 regions, including: three from Poland (out of 16) and Bulgaria (out of 6), one from Hungary and Romania, as well as Cyprus, Estonia, Lithuania, Latvia and also Malta.

The third class, covering 60 regions, (38 from EU15, i.e. 18,2% and 22 from EU 12, i.e. 39,3%) may be called a smart specialization class in industry, since it is characterized by the highest average share of workforce employed in industry (35,7%) and in high and mid-tech industry in the total number of workforce employed in industry (31,9%). The following regions are included in the third class: 21 from 39 German regions (which makes 1/3 of all regions in this class), 7 from 8 Czech, 6 from 21 Italian, 5 from 7 Hungarian, 4 from 22 French, 3 from 4 Slovak, 3 from 16 Polish, 3 from 17 Spanish; 2 Slovenian regions, 2 from 6 Bulgarian and 2 from 9 Austrian regions, as well as one region from Finland and one from Belgium.

The most numerous fourth class, covering 40,8% of all analyzed EU regions, i.e. 108, are characterized by the highest average share of workforce employed in services (75,4%) and also in knowledge-intensive services in the total number of workforce employed in services in the region (54,4%). This class may be called a set of regions following the path towards smart specialization in services. It covers 50,2% of regions from EU 15 and 5,4% of regions from EU 12 ones. Regions from the latest accessions represent three capital regions, or the ones including capital city: Praha, Közép-Magyarország and Bratislavský kraj. Individual regions from UE 15 (Luxembourg, Southern and Eastern, Comunidad de Madrid, Wien and 35 from 37 British regions (94,6%), 14 from 39 German (36%), 12 from 22 French (54,5%), 11 from The Netherlands (91,7%), all Swedish (8) and all Danish regions (5), 9 from 11 Belgian ones (82%), 4 from 5 Finnish (80%) and 3 from 21 Italian (14,3%) regions.

The obtained classes, however, included regions featuring 40% membership regarding the dominating, in a given class, property value and causing that further analyses focused only on these regions which met class membership threshold at the set value of 0,8.

The first class, which covers 16 regions after setting higher membership threshold, is characterized, just like in case of the initial division, by the highest mean value of X_1 property (26,02%) and the second mean value of X_2 property (28,8%) and also the lowest mean value of other properties, such as:

- workforce employed in services in the total number of workforce in the region (45,78%),
- workforce employed in knowledge-intensive services in the total number of workforce in services in the region (40,37%),
- workforce employed in high and mid-tech industry sectors in the total workforce number employed in industry (12,64%).

The first class covers 6% of the European space regions (5, i.e. 2,4% of EU 15 regions and 11, i.e. 19,6% of EU 12 regions) and includes regions from 4 countries: 7 Polish, 4 Romanian and 4 Greek regions and one from Portugal.

In the second class covering 42 EU regions (i.e. 15,8%) the second largest average levels of X_1 (6,7%) and X_3 (67,8%) properties and the third largest average levels of X_2 (25,7%), X_4 (41,6%) and X_5 (14,2%) properties should be registered. This class includes the following regions:

- 37 UE 15 regions (17,7%), including: 11 from 17 Spanish (64,7%), 8 from 21 Italian (38,1%), 5 from 9 Austrian (55,6%), 3 from 22 French (13,6%), 3 from 7 Portuguese (43%), 3 from 39 German (7,7%) and also one from Greece and one from Ireland,

- 5 UE 12 regions (8,9%): one from Bulgaria and Romania, as well as Cyprus, Lithuania and Latvia.

The third class, including 46 regions (28 from UE 15, i.e. 13,4% and 18 from UE 12, i.e. 32%), should be defined as the smart specialization class in industry (the highest average share of workforce in industry (36,84%) and in high and mid-tech

industry in the total number of workforce employed in industry (31,3%). This class includes the following regions: 16 out of 39 German (every third region in this class), 7 out of 8 Czech, 6 out of 21 Italian, 5 out of 7 Hungarian, 2 out of 22 French, 3 Slovakian and 3 Spanish, 2 Polish, one Slovenian and one Austrian region.

The fourth class is the biggest and covers 32,8% of the analyzed EU regions, i.e. 87 of them. It is distinguished by the highest average level of X_3 (75,9%) and X_4 (54,7%) properties, i.e. knowledge-intensive services. This class covers 40,2% regions from EU 15 and 5,4% of regions from EU 12 (capital regions or these including capital city: Praha, Közép-Magyarország and Bratislavský kraj). One region of UE 15 (Southern and Eastern, Comunidad de Madrid, Wien) and 32 from 37 British (86,5%), 9 from 39 German (23,1%), 11 from The Netherland (91,7%), 6 from 22 French (27,3%), 7 from 8 Swedish (87,5%), all Danish regions (5), 7 from 11 Belgian (63,6%), 2 from 5 Finnish (40%) and 2 from 21 Italian (9,5%) regions.

Regions which “left” the classification at the set membership threshold of 0,8 were mostly, following their initial assignment, included in the second class (difficult in terms of its clear and unique specification) – made up of 28 regions, including 5 from Greece (Ionia Nisia, Attiki, Voreio Aigaio, Notio Aigaio and Kriti), 4 from Italy (Provincia Autonoma Trento, Umbria, Abruzzo, Molise), 3 from Poland (Mazowieckie, Zachodniopomorskie i Lubuskie) and 3 from France (Champagne-Ardenne, Picardie, Languedoc-Roussillon), 2 from Bulgaria (Severozapaden, Severoiztochen), Spain (La Rioja, Aragón) and Portugal (Lisboa, Região Autónoma dos Açores), as well as one from Austria (Steiermark), Belgium (Prov. West-Vlaanderen), Germany (Weser-Ems), The Netherlands (Zeeland), Hungary (Dél-Alföld) and also Malta and Estonia. The fourth class (definitely service oriented one and characterized by high innovation level of services) covers 21 regions from such countries as: France (Centre, Pays de la Loire, Aquitaine, Rhône-Alpes, Auvergne, Provence-Alpes-Côte d’Azur), Germany (Oberbayern, Bremen, Mecklenburg-Vorpommern, Hannover, Saarland), Great Britain (Inner London, North Eastern Scotland, Northern Ireland), Finland (Pohjois-Suomi, Åland) and Belgium (Prov. Liège, Prov. Luxembourg), Italy (Valle d’Aosta), Sweden (Småland med öarna) and Luxembourg. The third class (mid and high-tech industry oriented) includes 14 regions from such countries as: Germany (Karlsruhe, Mittel-franken, Gießen, Braunschweig, Rheinhessen-Pfalz), Bulgaria (Severen tsentralen, Yugoiztochen), France (Lorraine, Alsace), Austria (Vorarlberg), Belgium (Prov. Limburg), Finland (Länsi-Suomi), Poland (Śląskie) and Slovenia (Vzhodna Slovenija). The first class (farming profile) – 11 regions from Poland (Wielkopolskie, Opolskie, Warmińsko-Mazurskie), Greece (Thessalia, Ipeiros, Dytiki Ellada) and Romania (Centru, Sud-Vest Oltenia, Vest), Portugal (Norte) and Bulgaria (Yuzhen tsentralen).

The membership threshold increasing strategy, in most cases, also resulted in decreasing extreme values of the analyzed properties (maximum and minimum) and the mean value in each class and, in consequence, in reducing standard deviation and variability coefficient values which resulted in higher cohesion of the obtained classes.

Table 4.

Number of regions from EU countries in classes obtained using fuzzy c -means method and for classes obtained at threshold membership value of 0,8

Country (number of regions)	Number of regions in a class and membership thresholds				Number of regions in a class and membership thresholds				Undefined
	1	2	3	4	1	2	3	4	
	0,409 -0,986	0,391 -0,999	0,392 -0,999	0,326 -0,999	0,800 -0,986	0,800 -0,997	0,800 -0,999	0,800 -0,999	<0,800
Austria (9)		6	2	1		5	1	1	2
Belgium (11)		1	1	9				7	4
Germany (39)		4	21	14		3	16	9	11
Denmark (5)				5				5	
Spain (17)		13	3	1		11	3	1	2
Finland (5)			1	4				2	3
France (22)		6	4	12		3	2	6	11
Greece (13)	7	6			4	1			8
Ireland (2)		1		1		1		1	
Italy (21)		12	6	3		8	6	2	5
Luxemburg (1)				1					1
The Netherlands (1)		1		11				11	1
Portugal (7)	2	5			1	3			3
Sweden (8)				8				7	1
Great Britain (37)		2		35		2		32	3
Bulgaria (6)	1	3	2			1			5
Cyprus (1)		1				1			
The Czech Republic (8)			7	1			7	1	
Estonia (1)		1							1
Hungary (7)		1	5	1			5	1	1
Lithuania (1)		1				1			
Latvia (1)		1				1			
Malta (1)		1							1
Poland (16)	10	3	3		7		2		7

cd. Table 4.

Romania (8)	7	1			4	1			3
Slovenia (2)			2				1		1
Slovakia (4)			3	1			3	1	
UE 27 (265)	27	70	60	108	16	42	46	87	74
UE 15 (209)	9	57	38	105	5	37	28	84	55
UE 12 (56)	18	13	22	3	11	5	18	3	19

Source: Author's compilation using **R** programme.

Abbreviations used in the study: UE 27 (regions of all UE countries), UE 12 (regions from countries of the latest two accessions), UE 15 (regions from "old EU 15"), AT (Austria), BE (Belgium), NL (The Netherlands), DE (Germany), DK (Denmark), ES (Spain), FI (Finland), FR (France), GR (Greece), IE (Ireland), IT (Italy), LU (Luxemburg), PT (Portugal), SE (Sweden), UK (Great Britain), BG (Bulgaria), CY (Cyprus), CZ (The Czech Republic), EE (Estonia), HU (Hungary), LT (Lithuania), LV (Latvia), MT (Malta), PL (Poland), RO (Romania), SI (Slovenia), SK (Slovakia).

6. FINAL CONCLUSIONS

The adequately focused regional policy is capable of releasing the EU growth potential by means of promoting innovation in all regions and, at the same time, ensuring complementarity between EU, national and regional support for innovation, research and development, entrepreneurship and also technology and communication. Regional policy represents the basic tool for adjusting EU innovation priorities for the purposes of practical, field activities.

Europe 2020 strategy covers three priority areas, five main goals, ten integrated guidelines and seven leading initiatives and also represents the development vision continuation, outlined by the Lisbon Strategy, as well as an attempt to prevent the slowdown of European economy growth, to counteract crisis effects which brought about the highest, in almost 80 years, economic downturn and unveiled extensive structural weaknesses of European economies.

The problem of R&D specialization is of particular significance for these regions and countries which do not play the role of leaders in each of the major domains in science and technology. Many regional specialists and politicians claim that these regions and countries need higher intensity of knowledge oriented investments in the form of high education and professional trainings quality, outlays on R&D (both public and private) and other activities related to innovation.

However, a question arises whether this presents a better alternative for the so far conducted policy, according to which even modest outlays on investments are necessary in many fields of technology, or in pioneering research (e.g. in biotechnology, information technology, nanotechnology), which unfortunately may not result in sig-

nificant effects in one particular area because of financial means dispersion. In view of the above observations, support and encouragement for investing in programmes, which may function as the supplement to assets held in a region or a country, seems a promising strategy aimed at creating future nationwide opportunities and establishing advantage in regional space.

As the result of smart specialization one should expect diversity among regions as the replacement of the system which creates more or less the same output in every region in a reproductive way, which also favours over extensive correlation and replication of R&D and educational investment programmes, which again reduces possibilities for supplementing each other within the framework of European knowledge base. Smart specialization represents both a concept and a tool supporting regions and countries in their response to the key question about their unique position in knowledge-based economy.

Research results have proved that we face a large number of regions which membership in a distinguished regional classes in Europe, regarding smart growth level, turns out difficult to define. Therefore it confirms that the approach to regions' classification, accepted in the hereby study, which consists in applying fuzzy classification methods is founded, since these methods provide much more additional information about classified regions than it case of classical methods.

Among regions of undefined membership, at the set threshold, 74 regions were listed (more than every fourth EU region), out of which 55 are from EU 15 (26,3%) and 19 from EU 12 (33,9%). The largest group covers German and French regions (11 each), and also Greek (8), Polish (7), Italian and Bulgarian (5 each), Belgian (4), Finnish, Portuguese, British and Romanian (3 each), 2 from Austria, 2 from Spain, and one from each of such countries as: The Netherlands, Sweden, Hungary, Slovenia, Luxemburg, Estonia and Malta. These regions may be defined as the ones "searching" for an optimum intelligent specialization path. They should become objects of particular attention and care for the individuals involved in managing development and responsible for allocating funds, as well as accountable for both interregional and intraregional policy. On the other hand, the regions classified in the third (mid and high-tech industry) and the fourth (knowledge-intensive services) class should be supported in their further development and activities have to be taken up in order to maintain their present development pace.

According to the authors the analysis of transformations dynamics and directions regarding regional membership to particular classes, as well as analyzing changes of such membership in particular EU countries, constitute an interesting focus for further studies.

LITERATURE

- [1] *A Digital Agenda for Europe*, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, EUROPEAN COMMISSION, COM(2010) 245 final/2, Brussels, 2010.
- [2] Bezdek J.C., [1981], *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York.
- [3] Bock H.-H., [2008], *Origins and extensions of the k-means algorithm in cluster analysis*, "Electronic Journal for History of Probability and Statistics", vol. 4, no. 2.
- [4] Cox E., [2005], *Fuzzy modeling and genetic algorithms for data mining and exploration*, Morgan Kaufmann Publishers, San Francisco.
- [5] *Digital Agenda Scoreboard*, Commission Staff Working Paper, SEC(2011) 708, European Commission, Brussels, 2011.
- [6] Domańska W., [2010], *Strategia rozwoju Europy do 2020 r.*, „Wiadomości Statystyczne” no. 8 (591).
- [7] Dunn A., [1973], *Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters*, "Journal of Cybernetics", vol. 3.
- [8] *EUROPA 2020. Strategia na rzecz inteligentnego i zrównoważonego rozwoju sprzyjającego włączeniu społecznemu*. Komisja Europejska, Komunikat Komisji, KOM(2010) 2020 wersja ostateczna, Bruksela 2010.
- [9] Giannitsis A., Kager M., [2009], *Technology and specialisation: Dilemmas, options and risks?*, Expert group "Knowledge for Growth".
- [10] Höppner F., [1999], *Fuzzy cluster analysis: methods for classification, data analysis, and image recognition*, John Wiley&Sons, Chichester.
- [11] *Innovation Union Competitiveness report*, Directorate-General for Research and Innovation, Directorate-General for Research and Innovation, Research and Innovation, European Commission, Luxembourg: Publications Office of the European Union, 2011.
- [12] Jajuga K., [1990], *Statystyczna teoria rozpoznawania obrazów*, PWN, Warszawa.
- [13] Lasek M., [2002], *Data Mining. Zastosowania w analizach i ocenach klientów bankowych*. Oficyna Wydawnicza „Zarządzanie i finanse”, Biblioteka Menadżera i Bankowca, Warszawa.
- [14] Leung Y., [1983], *Fuzzy Sets Approach to Spatial Analysis and Planning, a Nontechnical Evaluation*, "Geografiska Annaler. Series B", vol. 65, no. 2.
- [15] Nascimento S., Mirkin B., Moura-Pires F., [2000], *A fuzzy clustering model of data and fuzzy c-means*. IEEE International Conference on Fuzzy Systems: Soft Computing in the Information Age, vol. 1.
- [16] Pawełek B., [2008], *Metody normalizacji zmiennych w badaniach porównawczych złożonych zjawisk ekonomicznych*, Wydawnictwo UE w Krakowie, Kraków.
- [17] *Polityka regionalna jako czynnik przyczyniający się do inteligentnego rozwoju w ramach strategii Europa 2020*, Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów, KOM(2010) 553, Bruksela, 2010.
- [18] *Regions in the European Union. Nomenclature of territorial units for statistics NUTS 2006/EU-27*, Series: Methodologies and Working Papers, European Commission, Luxembourg 2007.
- [19] Sokołowski A., [1992], *Empiryczne testy istotności w taksonomii*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- [20] Steinley D., [2006], *K-means clustering: a half century synthesis*, „British Journal on Mathematical and Statistical Psychology”, vol. 59.
- [21] *Strategia na rzecz inteligentnego i zrównoważonego rozwoju sprzyjającego włączeniu społecznemu*, Krajowy program reform Europa 2020, Ministerstwo Gospodarki, Warszawa 2010.
- [22] Walesiak M., Gatnar E. (red.), [2009], *Statystyczna analiza danych z wykorzystaniem programu R*, Wydawnictwo Naukowe PWN, Warszawa.

- [23] Wintjes R., Hollanders H., [2010], *The regional impact of technological change in 2020 – Synthesis report*, European Commission, DG Regional Policy, Brussels.
- [24] Wysocki F., [2010], *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*, Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu, Poznań.
- [25] *Youth on the Move*, Publications Office of the European Union, European Union, Luxembourg, 2010.
- [26] Zimmermann J.H., [2001], *Fuzzy set theory and its applications. Fourth edition*, Kluwer Academic Publishers, Boston.

OCENA INTELIGENTNEGO ROZWOJU REGIONÓW EUROPEJSKICH Z ZASTOSOWANIEM KLASYFIKACJI ROZMYTEJ

Streszczenie

„Strategia Europa 2020” stanowi wizję rozwoju gospodarki europejskiej, dla której jednym z priorytetów jest rozwój inteligentny czyli oparty na wiedzy i innowacjach. Istotnym elementem inteligentnego rozwoju jest inteligentna specjalizacja obejmująca przedsiębiorstwa, ośrodki badawcze oraz szkoły wyższe, które współpracują na rzecz określenia najbardziej obiecujących obszarów specjalizacji w danym regionie. Stanowi ona zarówno koncepcję jak i narzędzie pozwalające regionom i krajom ocenić ich unikalną pozycję w gospodarce opartej na wiedzy. Trudno przecenić tą wiedzę na etapie formułowania założeń polityk regionalnych i interregionalnych oraz ustalania kierunków dystrybucji środków finansowych przeznaczonych na dalszy rozwój regionów budujących swoją przewagę w przestrzeni regionalnej oraz pozycję w gospodarce opartej na wiedzy. Dlatego zasadniczym celem niniejszego opracowania było wyodrębnienie klas regionów w przestrzeni europejskiej ze względu na zjawisko złożone jakim jest inteligentna specjalizacja. W tym celu zastosowano klasyczne i rozmyte metody klasyfikacji. Podejście takie umożliwiło m.in. wskazanie tych regionów, dla których nie można jednoznacznie określić przynależności do wyodrębnionych klas. Są to regiony „poszukujące” optymalnej ścieżki inteligentnego rozwoju, które winne zostać otoczone szczególną uwagą przez podmioty zarządzające rozwojem zarówno na szczeblu regionalnym, krajowym jak i całej wspólnoty europejskiej.

Słowa kluczowe: rozmyta klasyfikacja regionów, rozmyta metoda c -średnich, Europa 2020, rozwój inteligentny regionów

FUZZY CLASSIFICATION OF EUROPEAN REGIONS IN THE EVALUATION OF SMART GROWTH

Abstract

“Europe 2020 Strategy” presents the vision of European economy development, in which smart development, i.e. development based on knowledge and innovation, constitutes one of major priorities. Smart specialization which refers to enterprises, research centres and high schools cooperating in defining the most promising areas of specialization in a given region, represents one of crucial smart development

components. Smart specialization refers to both, the concept and the tool, allowing regions and countries to assess their unique position in knowledge-based economy. This knowledge should not be underestimated at the stage of preparing regional and interregional policy assumptions and specifying directions for the distribution of financial means allocated to further development of regions, constructing their advantage in regional space and position in knowledge based economy. Therefore, the essential objective of the hereby study is to distinguish classes of regions in European space with regard to one complex phenomenon, i.e. smart specialization. For this reason both classical and fuzzy classification methods were applied. Such approach facilitated e.g. specifying these regions for which it is difficult to provide clear division regarding their membership in distinguished classes. They are the regions which “keep searching” for their optimum path of smart development and which should be offered particular attention by entities managing development at regional, national and overall EU level.

Key words: fuzzy region classification, fuzzy *c*-means, Europe 2020, smart growth of regions