ARGUMENTA OECONOMICA
No 2 (37) 2016
PL ISSN 1233-5835

Małgorzata Markowska[*], *Marta Kusterka-Jefmańska*[**], *Bartłomiej Jefmański*[***]

ANALYSIS OF SMART SPECIALIZATION IN EUROPEAN REGIONS USING FUZZY CLASSIFICATION[1]

The major areas of specialization assessment cover for example export, R&D expenditure, employment rate, high-tech products (e.g. patents). Many measures indicated in the subject literature for specified areas cannot be currently determined at regional level due to data unavailability.

The article attempts to identify the areas and directions of changes in the smart specialization development of the European regions. An analytical approach, based on the application of fuzzy classification in the analysis of dynamic data, was used for this purpose. The analysis covered 265 European regions at NUTS 2 level, described by variables characterizing employment shares in economic sectors. The level of membership to the separated classes was estimated for each region in particular years of the adopted research period 2000–2010, which facilitated analyzing the directions of their location changes in the European area divided into classes. The obtained results allowed for the positive verification of the proposed hypotheses about the usefulness of fuzzy classification in the identification of areas and directions for smart specialization changes in regions establishing the European regional area. Both, the regions characterized by the changing sector system and the ones maintaining their sector specialization profiles in the adopted research period, were identified in the course of the conducted analyses.

Keywords: fuzzy classification, smart specialization, NUTS 2 regions

JEL Classifications: O18: Regional Analysis, C38: Cluster Analysis; O31: Innovation and Invention, O52: Europe

DOI: 10.15611/aoe.2016.2.02

1. INTRODUCTION

The concept of smart specialization appeared for the first time in EU strategic documents in 2010, in relation to the document: "Europe 2020. A strategy for smart, sustainable and inclusive growth" (Europe 2010a).

[*] Department of Regional Economy, Wrocław University of Economics
[**] Department of Quality Management and Environment management, Wrocław University of Economics
[***] Department of Econometrics and Informatics, Wrocław University of Economics

[1] The study was conducted within the framework of research grant National Science Centre no. 2011/01/B/HS4/04743 entitled: *European regional space classification in the perspective of smart growth concept – dynamic approach.*

Therefore it represents a relatively new idea. Smart specialization is related to one of three priority components covered by this strategy, namely to smart growth, i.e. a knowledge and innovation based economy (Regional 2010).

Defining smart specialization in relation to a region requires indicating selected areas (pre-selection) in order to focus on the identified regional resources while carrying out further strategic activities. Regardless of the economic stabilization level, or even its absence, if due activities are diversified, it is possible to identify certain key branches in a region.

The purpose of the article is to assess the usefulness of fuzzy classification in the process of identifying the areas and directions of smart specialization changes in regions constituting the European regional space at NUTS 2 level. The following research hypotheses were put forward for this purpose:

– it is possible to identify groups of regions characterized by a determined sector profile using the fuzzy classification of dynamic data,
– the suggested approach is a useful tool for the identification of the directions for smart specialization changes in the European regions.

The presented results constitute a continuation of the research conducted by the authors. Other studies present, among others, the classification of the EU regions at NUTS 2 level, but used a classic approach in determining the number of classes, whereas the analyses were performed based on data from one year (Markowska and Jefmański 2012a). The study also offers a detailed discussion of various variants for assigning regions to classes owing to the increasing thresholds. Further studies attempted to determine an initial research field and to define smart specialization areas in Polish regions based on a dynamic approach (Markowska and Jefmański 2012b).

2. SMART SPECIALIZATION – VISION, CONCEPTION, MEASUREMENT

2.1. The concept of smart specialization

Smart specialization represents an important justification of a certain policy and concept for the purposes of regional innovation policy. It propagates the efficient, effective and synergic implementation of public R&I (research and innovation) types of investment, as well as supporting member states and their regions in the process of the existing, leading sectors diversification and improvement (Malerba 2005, Malerba and Montobbio 2000), also in strengthening the inherent innovation capacity. In short, smart specialization consists in paying more attention to innovation and offering an innovation focused growth strategy which concentrates on the strengths and competitive advantage ingrained in regions. It refers to specializing in a smart manner, i.e. based on the obtained information and strategic intelligence in taking advantage of regional assets and skills aimed at learning which specializations may be both developed and implemented with reference to, for example those present in other regions (Foray and van Ark 2007).

The concept of smart specialization:
- indicates effective policy refinement by forecasting the creation of innovative and competitive units, clusters and/or regions,
- assumes intervention and therefore certain explicit and implicit tasks coupled with the purposeful concentration of resources in a certain form,
- establishes indispensable mechanisms of financial support offering the capacity to generate extensive, positive, external social effects in the future,
- assumes the presence of evaluation criteria regarding which specializations, and in consequence, which policy objectives represent the smart ones.

The idea of smart specialization should be referred to as the dynamic process of searching for proper areas to focus attention on. As such, smart specialization does not require forcing any specialization by means of any specific form of policy covering for example an overall industrial sector. On the contrary, it requires inventive entrepreneurial processes involving all stakeholders in order to identify and indicate what a given country or region is best at in the domain of science and technology, and also where to expect improvements. Such an inventive process has to be included among broader policy objectives and should also identify corporate governance mechanisms, as well as foster the criteria used as guidelines in making adequate choices.

Smart specialization requires strategic decisions and policy following an in-depth analysis. Priority-setting is performed based on the analysis of (Foray, David and Hall 2009):
- assets – industrial structures, clusters, universities, research, scientific, technical institutions, skills, human capital, environment, market access, management systems as well as relations and cooperation with other regions,
- challenges – e.g. population aging, imbalance in the job market, environmental aspects,
- competitive advantages and potential – SWOT analysis, forecasting and trends analysis, technology mapping, cluster analysis, and the knowledge of markets.

Smart specialization, having assumed that different paths of regional innovation and growth are possible, requires an integrated approach and a place-based approach for policy designing and implementation with operational principles adjusted to local conditions which cover the following[2]:

a) revitalization of traditional sectors by means of activities resulting in higher added value and seeking market niches;
b) modernization by means of new technologies approval and dissemination;
c) technological diversification from the existing specializations to related areas;
d) development of new business forms by means of radical technological changes and breakthrough innovations,
e) implementation of new forms of innovations, such as open and user-led innovation as well as social innovation and service innovation.

D. Ahner (2011, p. 2) states that: "Smart specialisation is definitely not about picking winners from above deciding where to place one or the other activity in Europe's map but, on the contrary, it is a place-based approach to identify opportunities for selective technological diversification in each region based on their own innovation capacities and economic and institutional structures. In other words, smart specialisation is about focusing on what you are good at and diversifying from your existing capacities into higher value added activities and emerging market opportunities in order to better position your cities and regions in the global value chains".

He emphasizes that smart specialization requires an iterative and dynamic approach which distinguishes the following phases (Comparative 2011):

– fostering and - where possible - channelling entrepreneurial experiments and discoveries,
– observing, monitoring and evaluating the progress of experiments,
– establishing and disseminating the strategic vision based on observation and assessment,
– identifying and addressing failures possibly preventing these experiments from becoming strong economic growth incentives (by offering training opportunities, developing skills and other stimuli for new activities enhancement).

D. Foray also defines smart specialization as an entrepreneurial process of discovery, identifying where a region can benefit from specializing in a particular area of science and technology[3].

---

[2] http://s3platform.jrc.ec.europa.eu

[3] Speech delivered at the conference on "Regions for Economic Change" on 23 June 2011.

2.2. Smart specialization in EU strategic documents

This embedded role of smart specialization in the Europe 2020 policy framework has been highlighted by the Council of the EU in its conclusions on the Innovation Union. The Council underlined “the concept of ‘smart specialization’, with each region building on its own strengths, to guide priority-setting in national and regional innovation strategies, as well as cross-border cooperation where appropriate” (Council 2010, p. 9) and invited the Commission “to advise Member States on possible improvement of the performance of their national innovation systems and with the implementation of smart specialization strategies” (Council 2010, p. 11).

Within the framework of the documents fostering the implementation of Europe 2020 strategy, the Innovation Union initiative (Europe 2010b) was, among others, prepared by the European Commission – it represents one of seven guideline initiatives targeted at the improvement of conditions and access to research financing which is supposed to facilitate situations when innovative ideas will be transformed into new products and services to foster economic growth and create new jobs (A strategy 2010). These tasks will be performed by strengthening the role of EU instruments which support this path of development including structural funds, rural areas development funds, schemes for competition support especially as the result of increasing research outlays and emphasis on smart specialization (Regional 2010).

Therefore the European Union funds exert the most important impact on the implementation of the Innovation Union priorities, as well as the practical activities within the framework of smart specialization in regions, since they constitute the financial instruments for cohesion policy, which is conditional and also concentrated more on output and efficiency. Therefore an *ex ante* conditionality, referring to smart specialization, will be introduced, which is supposed to imply extended expenditure on research, technological development and innovation. This is one of the eleven thematic objectives of the EU funds listed in the proposals defining common regulations for the structural funds and the cohesion fund (Annex 2011). Likewise, the same conditionality applies to theme one (Fostering knowledge transfer and innovation in agriculture, forestry and rural areas) of the European Agricultural Fund for Rural Development (EAFRD)[4].

Obtaining funds by a region is conditioned not only by the need to develop research and innovation strategy for the purposes of smart specialization (separating the part for digital development – extended access,

[4] http://ec.europa.eu/agriculture/cap-post-2013/legal-proposals/com627/627_en.pdf

implementation and quality of information and communication technologies), but also meeting the following premise by the member state:
- adopting national research and innovation strategy for smart specialization – based on a SWOT analysis and covering both a monitoring and review system – offering suggestions for research and development financing incentives by private capital,
- defining the level of R&D financing by the government,
- a long-term plan referring to the budget and priority investments related to EU strategic priorities,
- including chapters on digital development in the national policy framework for research and innovation regarding smart specialization, in relation to the thematic objective referring to broadband internet (as in regional strategies).

Defining smart specialization is the conditionality for accessing European funds in the new programming period, i.e. 2014–2020, as the component of new EU policy towards regions. Therefore regional authorities have an obligation of defining branch specialization, as well as specifying both the strengths and weaknesses in a region and also innovation based growth. Defining smart specialization should facilitate the concentration of resources on priorities of crucial significance (Dobrzycka 2012).

The support for countries in the process of smart specialization strategy construction is to be provided by the Smart Specialization Platform created in mid-2011 (Foray et al. 2012). It is highlighted that each and every region represents a particular individuality and therefore no ready-made solutions should be expected from the Platform, but rather suggestions of tools for creating them by offering ideas, facilitating contacts and exchanging opinions, experiences and knowledge as well as information between regions and the EU member states, the promotion of annual meetings involving politicians and environments responsible for smart specialization, and also preparing adequate methodology and a case studies base.

2.3. Types and measures of regional specialization

In literature, the measurement of specialization originates in trade theory. The first classical theory of specialization advantages is the one referring to comparative costs by D. Ricardo and represents the continuation of A. Smith's concept. A range of specialization indices were developed in order to reflect country specialization out of which various were used, after some adjustments, as technology specialization indices. The following examples can be listed (Giannitsis and Kager 2009):

- absolute export market share: share of the exported good of the country (region) in the total world (national) export of this good,
- relative export market share: measures the share of the exported good of the country (region) in the total world (national) export of this good to the equivalent of the trade partner's share in the world (national) export – defines the competitive position of a given country (region) in the world (national) market in relation to a competitor's position,
- export specialization indicator: measures the relation of goods export of the country (region) to the total export of this country (region) divided by the relation of goods from another geographical area (chosen partner country (region), OECD, EU) – captures comparative advantages, since it compares exports from two analyzed areas, e.g. the country and the OECD area of the certain good to their total export understood as the collective, alternative good,
- indicator of the disclosed comparative advantages (the Balassa index): illustrates combined country competitive and comparative advantage – share of goods export from the country (region) in the total export of this good to the share of the total country (region) export in the total world export,
- adjusted index of revealed comparative advantages: simplified version of the Balassa index – estimated as difference of shares of a given good in the total export and import of the country without making comparisons to the adequate proportions in partner countries worldwide.

While analyzing the technological advantages or country (regional) specialization, instead of data referring to export, those considered are, for example (van Zeebrock et al. 2006; Alcorta, Peres 1998): data referring to selected economy sectors (e.g. R&D, Larosse and Pontkakis 2008), high-tech products, employment structure, added value or specific variables (patents). The other, above presented, indices may also be adjusted to measure the changes of technology specialization over time.

All these indices measure whether, and to what extent, the studied country, region, firm or supranational area presents a higher concentration of the chosen variable in these areas where they possess higher specialization than other competitors. This empirical observation does not cover the underlying reasons and the process leading to such a specialization. It also does not refer to the influence of the public, private or mixed nature of public resources including the role of institutions and other factors. However, from the perspective of policy, it is important to consider how specialization positions could be accomplished and what specific economic or social impact they may have (Giannitsis and Kager 2009).

Traditionally, as indicates Cutrini (2006), the absolute indices most frequently used in regional studies are the Gini coefficient, the Herfindhal index (e.g. in Sapir 1996, Aiginger and Pfaffermayr 2004) and the entropy index (used by Aiginger and Pfaffermayr 2004, Aiginger and Davies 2004). The Gini coefficient, however, has the disadvantage of placing higher importance on changes in the middle intervals of the income distribution (see Cowell 1995).

Cutrini (2006) emphasizes that the Gini location quotient, based on the Hoover-Balassa Index[5], is the most commonly used index for measuring relative concentration and relative specialization in many country-based empirical studies (Kim 1995, Amiti 1999, Haaland et al. 1999, Brülhart and Torstensson 1996, Brülhart 1998, 2001).

Additionally, the relative specialization based on the Manhattan distance (used, among others, in Krugman 1991, Hallet 2000, Midelfart et al. 2004, Mulligan and Schmidt 2005) presents each of the relations between numerator and denominator of the Balassa Index as a difference rather than a ratio, and is obtained through summing up such differences across all sectors in the case of specialization and across all the regions in the case of concentration (Cutrini 2006).

It is commonly known that the broad localization concept has two connotations, namely, regional specialization and industrial concentration.

The following indicators are also applied in the assessment of regional specialization:

- the Krugman index applicable in the assessment of a given region economic structure in relation to the average economic structure of other regions (Doświadczenia 2008); the most frequently compared variables are: employment share in a given sector in a region and average employment in a country, as well as the share of gross added value generated by a given sector in GDP of a region and average added value in this sector in the country,
- the P. S. Florence local specialization index (Florence 1939, 1944) represents a relative measure and illustrates the relation of a given quality percentage share in the structure of the *i*-th spatial unit or object to the percentage share of a given quality in the structure of a hierarchically higher unit.

Regional specialization may be evaluated using the methodology of clusters strength assessment applied by the European Cluster Observatory (Ketels and Sölvell 2006). The European Cluster Observatory provides a

[5] A different version of the Gini coefficient location is used in Midelfart et al. (2004).

unique cluster mapping and benchmarking methodology which is currently applied to more than 35 manufacturing and service sectors in 32 EU and EU-associated countries.

Both economic situation and attractiveness are closely related to regional economy profile, i.e. this specialization level. The problems of regional economy condition and its future also refer to the identification of the specialization level, the implementation of resources available in a region, the number and condition of economic entities and/or the development of key branches. On the one hand, a specialization is the source of comparative advantages, the basis for competitiveness and the core concept of clusters and, on the other, it results in higher sensitivity to external shocks and the risk of adverse changes multiplication, which could be triggered by negative events (the domino effect). Therefore the relationship between specialization level and economic development dynamics cannot be referred to as a positive correlation resulting from necessity. Regional specialization is defined by comparing the regional economic structure to the average structure of other regions. High regional specialization occurs when a region has a significantly different branch structure than the average one characteristic for other regions (Specjalizacja 2008).

Regional specializations result from diverse reasons and mainly from the availability of natural resources (e.g. mining and excavation, forestry products), workforce resources, qualified workers, costs, infrastructure, legislation, climatic and topographical conditions (especially these related to tourism-oriented activities) and markets' proximity. Local and regional specialization are the derivatives of an overall investment attractiveness. Types of activities characterized by specific location factors constitute an exception in this area (Godlewska-Majkowska 2009).

The most common application of the "specialization" concept refers to the different significance of productive activities in the production structure of a country. More specifically, it is the technology specialization (and specialization in general) which constitutes the relative measure and may be defined by means of two different comparisons (Giannitsis and Kager 2009):

– the comparison between the relative weight of the reference variable (scientific knowledge, research, technology outputs, patents and/or productive areas) in the same country (region), e.g. specialization in ICT, biotech, electrical engineering etc.,
– the comparison between, respectively, the above national (regional) technology specialization patterns and similar figures of the third countries (regions) or areas.

– Specialization, by definition, illustrates two contradictory aspects (McCann and Ortega-Argilés 2001):
– a positive aspect, by means of indicating areas in which a country, region, sector and/or company presents a stronger position than other countries, sectors, regions and/or companies,
– a negative aspect, by pointing to areas of relative weaknesses.

In consequence, specialization in its positive sense, inherently implies the *concentration of capabilities* in some areas of knowledge, while in its negative sense it indicates weak capabilities in other areas when referred to the benchmark. The underlying concept of specialization, however, rejects the option that a country (region) achieves specialization positions in the overall and extensive spectrum of technologies (Giannitsis and Kager 2009).

Giannitsis and Kager (2009), indicate that specialization has various impacts depending on the technological level it refers to, whereas the scope of taxonomy can be as follows:
– scientific knowledge specialization,
– specialization in technology and innovation,
– specialization related to production processes,
– specialization related to clusters,
– horizontal and vertical specialization.

Following the above presented discussion on smart specialization, one may expect that specialization, as a policy concept, may become the catalyst of numerous ideas and initiatives undertaken by the European Commission and the OECD[6]. However, in order to capture politicians' attention and transfer the discussion from problems of a conceptual nature to research and empirical evidence, it is necessary to indicate the feasibility of statistical measurement which should encourage the authorities (at national and regional level) and experts to participate in the project (David, Foray and Hall 2009).

David, Foray and Hall (2009) emphasized that once deprived of measurements resulting in the specification of indicators, as well as the regular collection of statistical data, smart specialization would present just an idea for which monitoring progress, structural transformations assessment and performing efficiency comparisons would prove impracticable. S&T data and indicators are indispensable for the due condition visualisation,

[6] There is a reference to "smart specialization" in the Communication on the EU2020 Strategy (page 7); many linkages are also present between "smart specialization" and several themes of the work programme 2011–2012 of the Committee for S&T Policy at the OECD.

monitoring and evaluation of the designed and carried out innovative, for science and technology, solutions (David, Foray and Hall 2009).

At the current stage of study, conducting pilot research is necessary based on the existing statistical data, in order to confirm that smart specialization may be presented in a measurable manner and that aggregate measures specification is possible for this concept measuring. An initial indicator framework became the starting point and has been subject to subsequent discussions, modifications and will also be tested by different workgroups bringing together the interested individuals (Foray 2011).

In many regions smart specialization is closely related to a specific sector (e.g. the paper production industry, fishery, tourism). Therefore in the case of the already identified areas, the question should be asked about patents in this area worldwide in order to find out which technological activities are undertaken in this field by companies representing the same sector in other regions. This is, obviously, a mimetic approach, but since the activity is new enough it is possible that followers would not lag behind by very much (Foray, David and Hall 2009).

Studies focused on indicators take two problems into consideration: the discovery process and the tracking progress. The suggested, initial areas of assessment cover the following issues (David, Foray and Hall 2009):

– the process of entrepreneurial discovery – the intensity of companies' undertaking new, non-traditional activities,
– simple technology specialization indicators using patents, tracked over time to trace their evolution,
– innovation representing GPT (General Purpose Technologies) applications co-invention in the main local economy sectors, such as patent data or backward citations usage,
– correlation between R&D and training specialization and also economic structures as the guidelines for future budget allocations,
– smart specialization networks between regions: co-patenting, co-publication,
– clusters representing the geographical manifestation of smart specialization.

3. CLUSTER ANALYSIS METHODS IN REGIONAL STUDIES

Cluster analysis methods were described in detail, among others, in the studies by: Hartigan (1975), Aldenderfer and Blashfield (1984), Kaufman and Rousseeuw (1990), Gordon (1999), Everitt *et al.* (2011). The application of cluster analysis in regional studies allows for distinguishing regions characterized by similar levels of the analyzed economic phenomenon, or featuring similar economic structures, and combining them in homogenous groups as well as separating the development phases of the studied regions. Regional classification results are usually applied for the purposes of a typological identification of regions and for the assessment of their structural transformations (Strahl 2006).

In regional studies, cluster analysis methods are frequently applied in combination with other multidimensional statistical methods. Cluster analysis and discrimination analysis were used in the study by Hill *et al.* (1998), where one of the objectives was to distinguish clusters of major cities across the USA. Such an approach was also applied in the study by Stimson *et al.* (2001), in which hierarchical cluster analysis was implemented for identifying nine clusters of community opportunity and community vulnerability across Australia's metropolitan regions. In the case of both papers the application of a discrimination analysis was carried out in order to identify the most differentiating variables referring to the obtained classes. In the study by Mazzocchi and Montresor (2000), an assessment of farming development dynamics in 341 communes representing one of Italian regions, was performed using both the *c*-means method and the analysis of key components. The authors of the study, edited by Zeliaś (2000), having applied the Ward's and k-means methods prepared the assessment of spatial diversification referring to the population's living standard in Poland. Their work resulted, among others, in distinguishing typological groups of Polish regions in the period taken for analysis, and also in the classification results review in a dynamic perspective. The objective of the study by Pettersson (2001), was to support planning processes at regional level and for that purpose 500 micro-regions were classified in seven classes. The classification, applying the Ward's method, was performed based on indicators referring to the general goals of Swedish regional policy. The paper by Soares *et al.* (2003) discussed the new approach to Portuguese regions' classification in relation to their socio-economic growth, which was focused on supporting regional development policy. The study combined the application of a factor analysis and cluster analysis (using the Ward's and k-means methods) as a result of which three classes of regions were distinguished and characterized. In a similar way and for the same reason, Slovenian regions were classified (Rovan and Sambt 2003). The Ward's method and k-means method were also used in the study by del Campo *et al.* (2006), where European regional space was divided into four classes with regard to their socio-economic growth. The analysis presented by Cziráky *et al.* (2006), applies the combination of structural equation models and cluster analyses for an assessment of the regional development level in Slovenia and Croatia. Based on the Ward's method and k-means method, three groups of regions were distinguished in each of the two analyzed countries. The combination of the cluster analysis method (the Ward's method, the complete linkage method, the k-means method) and also Fisher's linear discrimination function constituted the grounds of the dynamic classification algorithm (suggested by Gertsbakh and Yatskiv 2006), applied in order to distinguish groups of the European Union countries regarding their economic growth. Many suggestions related to the possibilities for cluster analysis application in regional research are also presented in the study edited by Strahl (2006).

3.1. Fuzzy cluster analysis

Fuzzy cluster analysis represents a relatively new approach to the problem of regional classification. In a classic cluster analysis, a region's membership to a class is expressed by a binary variable. In other words, either a region belongs to a given class or it does not. The application of the fuzzy sets concept to the problem of classification allows for the option of a membership to more than just one class. This is possible due to binary variable substitution by a continuous variable taking values from the interval [0; 1]. This procedure allows for describing the situation more precisely, i.e. where the boundaries between classes are "unclear" and assigning a region to a unique class becomes more difficult (Jajuga 1990). 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 the approach where regions are assigned to one class only.

One of the first studies to apply fuzzy cluster methods in regional research was the article by Harris *et al.* (1993). The objective of the paper was to specify the typology of Nevada counties. The estimated membership degrees allowed for indicating specific economic development programmes for aggregate and individual counties. In the study by Topaloglou *et al.* (2005), the fuzzy *c*-means method was applied to present the typology of European regions situated along national borders (at NUTS 3 level) and also to assess each region's relative position in the European regional space. The fuzzy *c*-means method was also used in the paper by Nowak (2008), the main aim of which was to analyze the diversity level of 45 Polish sub-regions with regard to their economic activity level characterized by means of socio-economic variables. Manca and Curtin (2012), having applied the fuzzy *c*-means method, distinguished clusters of regions in the area of the Sardinia geopark for the purposes of regional development planning. Economic, social and geographical variables constituted the basis for the conducted classification.

3.2. Fuzzy *c*-means method

The fuzzy *c*-means method was proposed by Dunn (1973), and later generalized by Bezdek (1981) and Höppner (1999). It is an iteration method, the idea of which is very close to the classic k-means method and focuses on finding such gravity centres of classes which minimize the function below (Bezdek 1981):

$$J_m = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^m \left\| \mathbf{x}_j - \mathbf{v}_i \right\|^2 , \tag{1}$$

if:

$$\sum_{i=1}^{c} \mu_{ij} = 1, \ (j = 1,2,...,n) , \tag{2}$$

$$\sum_{i=1}^{n} \mu_{ij} > 0, \ (i = 1,2,...,c) , \tag{3}$$

$$\mu_{ij} \geq 0, \ (i = 1,2,...,c; j = 1,2,...,n) , \tag{4}$$

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

$\left\| \mathbf{x}_j - \mathbf{v}_i \right\|$ – the Euclidean distance between the *i*-th fuzzy class gravity centre $\mathbf{v}_i$ and the *j*-th object $\mathbf{x}_i$,

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

The fuzzy *c*-means method requires three parameters to be *a priori* defined: *m* – the fuzzy parameter, the number of fuzzy classes and the initial degree of membership in a class (or initial gravity centres of classes).

The parameter *m* specifies the degree of classification results as fuzzification. Its values should be greater than 1 where values close to unity bring about results similar to those obtained by means of classic methods. An increase in the value of parameter *m* results in the fact that the degree of objects membership in particular classes will take values close to the inverse number of classes, i.e. *1/c* (Lasek 2002). The literature does not offer a theoretical background for the choice of the optimum value of the parameter *m*, therefore its choice is frequently made based on experience coming from previously conducted empirical studies. Bezdek *et al.* (1984) indicate that the parameter value should belong to the interval [1,5;3]. Wysocki's research results suggest that the parameter value should be included in the interval [1,3;1,5] (Wysocki 2010).

Another parameter to be defined by the researcher is the decision regarding the optimal number of classes. Literature most frequently suggests two approaches in this matter (Wysocki 2010):

– 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 the fuzzy classification quality index reaches an extreme level.

The initial classification of objects can be done based on (Wysocki 2010):

– randomized assigning objects to classes,
– classification results obtained by applying disjoint classification for the same set of data,
– non-statistical knowledge.

4. THE CLASSIFICATION OF EU REGIONS BY THE LEVEL OF SMART SPECIALIZATION

4.1. Basic statistics of the analyzed variables

As has already been mentioned in the introduction to smart specialization identification, the preliminary selection of areas turns out to be indispensable, which allows for the concentration of strategic activities on specific regional resources. Regardless of the economic specialization level characteristic for a given region, or even the absence of such specialization (the activities are highly diversified) the specification of key branches is possible.

The selection of characteristics for the identification of areas investigated for the smart specialization of the EU regions resulted from the following, co-occurring reasons: 1) subject literature review regarding the areas of specialization assessment and measures; 2) the continuation of a "pillar" approach to smart development measurement (innovation, creative regions, smart specialization) within which similar measures were indicated for the identification of smart specialization (Markowska and Strahl 2012);

3) availability of comparable and relatively complete statistical data at regional level. Moreover, it seems that one of the first specialization symptoms is the growing importance of a particular economy sector as the background of the remaining ones. Therefore, as the indicators facilitating the identification of smart specialization areas, the importance – employment shares – of particular economic sectors in the entire economy was indicated.

In this paper the research set of objects is made up of NUTS 2 level regions, however, due to the absence of full data referring to the selected characteristics regarding French overseas regions (Guadeloupe, Martinique, Guyane, Réunion) and two Spanish regions (Ciudad Autónoma de Ceuta, Ciudad Autónoma de Melilla), 265 out of 271 regions (Regions 2007) were analyzed, i.e. 97.8% of all the distinguished ones.

Data regarding the selected characteristics originate from the Eurostat base and cover the period of 2000–2010. The following characteristics were suggested to define the preliminary investigation area focused on regional smart specialization:

AGRI – employment share in the farming sector in the total number of employment in the region,

MANUF – employment share in the industry sector in the total number of employment in the region,

SERV – employment share in the services sector in the total number of employment in the region,

KIS 2 – employment share in the knowledge-intensive services in the total number of employment in services in the region,

HIT 2 – employment share in the high and mid-tech industry in the total number of employment in industry in the region.

The data referring to employment originate from the Eurostat[7] database and cover those aged 15–64. All values are presented as a percentages and NACE Rev. 2 (NACE 2008) classification was used to define both the economy sectors and the other variables (KIS 2 and HIT 2). The diversification of characteristics applied for the purposes of preliminary selection referring to the smart specialization investigation area in the EU regions in the period of 2000–2010 is illustrated in Table 1.

Having analyzed the maximum values for the five studied qualities it was observed that only in the case of SERV an increase in value was registered in 2010 compared to 2000, while for the other qualities only lower values were obtained. In relation to the lowest values of the studied qualities the

[7] http://ec.europa.eu/eurostat/statistics-explained/index.php/Employment_statistics

situation was as follows: in case of AGRI in the period of 2000–2010, slight fluctuations were noticed (in the range of 0.19%–0.49%), however the values were identical in boundary years. A drop in the minimum value was registered regarding MANUF (from 12.3% to 7.2%), while for the rest of the qualities the minimum value presented an increasing tendency (the most in the case of SERV – from 19.4% up to 30% and also in the case of KIS 2 – from 23.8% up to 36.5%).

Very similar mean values and median values are characteristic for MANUF, in SERV and also HIT 2 and KIS 2 qualities, but a drop in median value and mean value was observed for three qualities (AGRI, MANUF and also in HIT 2), while an increase was registered for two qualities (SERV and KIS 2).

Standard deviation, which illustrates absolute diversification regarding the analyzed qualities, shows the range of 6.97–7.13 in the case of MANUF, up to 11.03–9.93 for SERV, however for the majority of qualities its values present a decreasing tendency (except for MANUF).

The highest variability (assessed as a percentage relation of standard deviation and mean value) was shown by AGRI (always over 112%), next HIT 2 (41.3%–47%), while the lowest and decreasing variability was observed for KIS 2 (17.6%–11.8%).

4.2. Fuzzy classification of smart specialization European regional space

Therefore the paper provides unified values for all variables by means of a standardization formula application. The Euclidean distance was used for the calculation of objects similarity. The value of the fuzzy parameter was defined at $m = 1.5$ level. The application of the fuzzy *c*-means method also requires settling the key issue, i.e. the number of classes. In the case of the absence of non-statistical information, the number of classes can be specified based on the selected disjoint classification method, or by performing a fuzzy classification procedure for a different number of classes, and an optimal number of classes can be decided based on the chosen measure for the estimation of classification quality (Wysocki 2010, p. 136). This paper follows the second approach. Literature presents numerous proposals on this matter. The review of measures regarding fuzzy classification quality estimation for the fuzzy *c*-means method is included in, for example, the study by Wang and Zhang (2007). Due to the fact that it is difficult to specify such a measure which, owing to its attributes, would excel over the others, it was decided that for the purposes of this study a few of them will be estimated, and based on their indications an optimal number of classes will be defined. The following measures are used in the paper: *PC* (*Partition Coefficient*), *PE* (*Partition Entropy*) (Bezdek 1974), *XB* (*Xie-Beni Index*) (Xie and Beni 1991), *FS* (Fukuyama and Sugeno 1989) (see Table 2).

In the case of the first two indices, their modified values were estimated following formulas presented respectively in Wu and Yang (2005), as well as Masson and Denoeux (2009). The initial version of the measures assumes a monotonous increase of their values along with the larger number of classes. On the other hand, the application of *XB* and *FS* measures allowed for considering, apart from the information about membership degree, also some aspects of the data set geometrical structure. As far as the problem of the optimum number of classes selection is concerned, in the case of the *PC* measure it is suggested by its maximum value. In relation to other measures, an optimum number of classes is indicated by their minimum values. The obtained results suggest, in most cases, the division of regions into four classes and therefore it is this measure which becomes fundamental in subsequent analyses presented in this paper.

The EU regions classification was performed based on one data set combining data from all the analyzed years. This means that each EU region was presented 11 times as an object (operational taxonomic unit). Therefore it constituted a separate object-period unit in every moment of the analysis. In consequence, having analyzed an 11-year-long period and having used 265 European regions for analysis, 2915 object-period units became the basis of classification. As a result, classes of regions were defined only once and their composition did not change throughout the overall period accepted for the research. Such an approach allows for the identification of regions which were always characterized by a certain smart specialization profile and also those which changed profile or underwent the process of changing it. In order to define both the directions and dynamics of these changes the degrees of these object-period units membership were used in the distinguished classes.

4.2.1. Interpretation of clusters

Four fuzzy classes can initially be characterized with regard to mean values of each quality in the distinguished classes (see Table 3):

– the first class includes regions characterized by the highest, out of all classes, mean value of AGRI, the third consecutive mean value of MANUF and the lowest mean values in the case of the other three qualities (for the purposes of this paper let us call this class a farming one),

- the second class is characterized by the highest SERV and in KIS 2, as well as the second highest mean value of HIT 2 (knowledge-intensive services class),
- the third class is characterized by the second highest mean value of workforce share in each of the three sectors (farming, industry and services) and the third highest mean value of HIT 2 and KIS 2 (poly-functional class),
- the fourth class covers regions featuring the highest of all the classes mean value for MANUF and also high and mid-tech industry (high and mid-tech industrial class).

Further evaluation of the obtained classes results in the following observations (see Table 4):

- the first class included two operational taxonomic units with a membership value below 0.4, 17 units presented a membership degree range of 0.4–0.5, 11 the range of 0.5–0.6, 19 the range of 0.6–0.7 and 22 where the analyzed value ranged between 0.7 and 0.8. The range exceeding 0.8 was characteristic for 149 units included in the first class;
- the second class covered four operational taxonomic units featuring a membership degree below 0.4, a membership degree ranging 0.4–0.5 was registered for 34 object-period units and the other ranges of membership degree were as follows: 83 units – 0.5–0.6 range, 88 units – 0.6–0.7 range, 98 units – 0.7–0.8 range and a range exceeding 0.8 was observed in the case of 837 object-period units;
- the lowest membership threshold (below 0.4) was measured for 10 operational taxonomic units in the third class, while in relation to particular value ranges the situations of object-period units positioning was as follows: 0.4–0.5 (64), 0.5–0.6 (115), 0.6–0.7 (111), 0.7–0.8 (138) and 357 presenting a membership degree to the third class higher than 0.8;
- the fourth class covered 9 operational taxonomic units with a low membership value (below 0.4), in the case of 37 units the degree of membership was ranging between 0.4 and 0.5 and respectively: 62 units – 0.5–0.6 range, 75 units – 0.6–0.7 range, 107 units – 0.7–0.8 range and for the largest number of object-period units (465) the degree of membership to the fourth class exceeded the value of 0.8.

4.2.2. EU regions in dynamic fuzzy classification

The composition of particular classes regarding the regions constituting them, was analyzed jointly in the period of 2000–2010, as follows (see Table 4):

- the first class registered 220 object-period units (regions from six countries: Greece, Romania, Poland, Portugal, Bulgaria and Spain) and after defining the threshold at the 0.8 level their number dropped to 149, i.e. 76.7% of all those studied in this class (from four countries: Greece, Romania, Poland and Portugal),
- the second class presented a total of 1144 objects (regions from 21 EU countries except regions from Bulgaria, Poland and also Cyprus, Estonia, Lithuania and Latvia) and after defining the threshold at the level of 0.8 the number of object-period units shrank to 837, i.e. 73.2% of all the objects analyzed in this class (from 18 EU countries, apart from the already listed ones and the regions from Greece, Romania and also Luxembourg),
- the third class covered 796 object-period units in a dynamic classification (from 22 countries excluding Danish, Finnish, Dutch, Swedish and Slovenian regions) of which 357 were left after introducing the membership threshold, i.e. 45% of them (regions from 15 countries, apart from the above listed regions and also Belgian, Czech, Slovak regions, as well as Luxembourg and Malta),
- the fourth class was made up of 755 object-period units (regions from 19 countries, excluding Danish, Greek, Portuguese regions and also Lithuania, Latvia, Luxembourg, Estonia and Cyprus), after defining the threshold at the level of 0.8 the number of object-period units was cut down to 465 (i.e. 76.7% of all the regions studied in this class). These regions originated from 14 countries with 42% of object-period units in this class coming from Germany).

Having defined the membership threshold, out of the 2915 object-period units, 62.1% of all the object-period units were left in the classes.

Among the countries featuring the highest share of regions which, owing to the lower than 0.8 membership threshold, could be referred to as "undefined" the following are included: Greece, Portugal and Poland (over 60% object-period units each) and out of the NUTS 1 level regions – Malta, Lithuania and Luxembourg.

Regions characterized by a strong economic profile from an employment point of view, are the regions which during the past 11 years maintained a similar employment structure which resulted in their classification, as object-period units, always in the same class. There were 156 such regions, i.e. 58.9% out of 265 analyzed EU regions. Their positioning in the classes was as follows: 13 regions in the first class, 72 in the second, 35 in the third while the fourth class always included 36 regions (see Table 5).

Particular classes, after reducing the number of units they included into those featuring membership degree equal or higher than 0.8, are characterized below.

The smallest group (the first class) covers regions characterized by definitely higher than in the case of other classes' share of AGRI. This class includes six Romanian, four Polish and three Greek regions (see Table 4).

The first class presents almost four times higher employment in farming than the mean value for the EU regions. The mean workforce share in industry is lower by two percentage points, as well as the average workforce share in services is lower by over 21 percentage points than the EU average. The means regarding HIT 2 and KIS 2 are also lower by nine percentage points.

The assessment of regions regarding distance from a class gravity centre, i.e. in relation to the similarity level of their global mean values and the analyzed characteristics with mean values in this class (based on the Bray and Curtis (1957) measure), illustrates that the most typical region in the class was Świętokrzyskie, for which the Bray and Curtis measure, standardized in the range [0;1], was the lowest (0.03).

In the case of these regions, smart specialization should be sought for mainly in farming sector sections and in research focused on its strengthening, increased competition and seeking innovative solutions which facilitate such activities (Proposal 2011, Regional 2010).

The second class is characterized by definitely the highest, comparing to the other classes, SERV and also in high-tech services, as well as the second highest HIT 2 mean value and the lowest mean value related to the workforce share in farming and industry sectors.

Out of 72 regions which were always, i.e. 11 times, included in the second class the following can be listed: British (23 regions), Dutch (11), Belgian and Swedish (7 regions each), Danish (5), German and Finnish (4 regions each), French (3), Italian (2) and also one Czech, Spanish, Slovak, Hungarian, Austrian and Irish region. This class covered 16 capital regions or regions including the country capital (see Table 5).

While assessing mean group values in relation to EU mean values regarding the analyzed qualities, it was observed that regions included in this class were characterized by two times lower average AGRI, MANUF lower by 5 percentage points and a higher average SERV (by 9 percentage points), in knowledge-intensive services (by 7 percentage points) and high-tech industry (by 2 percentage points).

The East Anglia region was the most typical in this class with a Bray and Curtis measure equal 0.006 which means that the global mean values characteristic for this region were the closest to the mean values for each variable in this class.

Broadly understood services should represent a smart specialization investigation area for regions included in this class and especially these which Eurostat defines as knowledge-intensive services (KIS) – NACE Rev. 2 divisions are given in parentheses(NACE 2008):

- knowledge-intensive high-tech services:
  - Post and Telecommunications (64),
  - Computer and related activities (72),
  - Research and development (73),
- knowledge-intensive market services (excluding financial intermediation and high-tech services):
  - Water transport (61),
  - Air transport (62),
  - Real estate activities (70),
  - Renting of machinery and equipment without operator, and of personal and household goods (71),
  - Other business activities (74),
- knowledge-intensive financial services:
  - Financial intermediation, except insurance and pension funding (65),
  - Insurance and pension funding, except compulsory social security (66),
  - Activities auxiliary to financial intermediation (67),
- other knowledge-intensive services:
  - Education (80),
  - Health and social work (85),
  - Recreational, cultural and sporting activities (92).

The third class regions may be referred to as poly-functional. They represent regions featuring the second highest average workforce share in each of the three sectors: farming, industry and services and the third highest mean value for KIS 2 and HIT 2 qualities. This class, in each of the analyzed years, was made up of 35 regions, including: Spanish regions (10), Italian (7), Portuguese (4), Bulgarian and Polish (3 regions each), Greek (2) and one Austrian and Hungarian region, as well as Estonia, Cyprus, Lithuania and Latvia.

While comparing the mean values of qualities in the class to those set for all EU regions it has to be indicated that in class three they are the closest. Castilla y León was the most typical one in the class for which the Bray and Curtis measure was 0.019.

It is difficult to define the area in which smart specialization range should be defined since this class represents quite a broad spectrum of the studied variables.

In the fourth class the mean values of qualities chosen for the preliminary selection of regions, in order to define their smart specialization, determine assigning regions to the high-tech industry group since it is the average MANUF and high and mid-tech industry which is the lowest in these regions. Additionally, these regions are characterized by lower than EU average AGRI, SERV and KIS 2.

The fourth class covers the following 36 regions: German (14), Italian (6), Czech (5), Spanish and Hungarian (3 regions each), Austrian (2), and one French, Slovak and Slovenian region. The Italian Emilia-Romagna region was the closest to the average values regarding the mean values of all its characteristics – the Bray and Curtis measure value for this region was 0.01.

In the case of this class, a regions smart specialization should be sought in the industry sector and especially high and medium high-technology manufacturing (in brackets are the NACE Rev. 2 divisions; NACE 2008):

- High-technology:
  - Manufacture of pharmaceuticals, medicinal chemicals and botanical products (24.4),
  - Manufacture of office machinery and computers (30),
  - Manufacture of radio, television and communication equipment and apparatus (32),
  - Manufacture of medical, precision and optical instruments, watches and clocks (33),
  - Manufacture of aircraft and spacecraft (35.3),
- Medium-high-technology:
  - Manufacture of chemicals and chemical products (24, excluding 24.4),
  - Manufacture of machinery and equipment n.e.c.[8] (29),
  - Manufacture of electrical machinery and apparatus n.e.c. (31),
  - Manufacture of motor vehicles, trailers and semi-trailers (34),
  - Manufacture of other transport equipment (35, excluding 35.1 Building and repairing of ships and boats and excluding 35.3).

[8] Not elsewhere classified.

The results of the performed classification, illustrated by the attached maps (see Figures 1 and 2), indicate that in many countries, knowledge-intensive services at regional level in 2000 are characteristic especially for capital territories or including the country capital, e.g. Spain, France, Germany, the Czech Republic, Slovakia, Austria and Denmark (see Figure 1).

According to the conducted research, 2010 is characterized by the definite predominance of regions featuring above-average SERV and mainly knowledge-intensive ones, especially in Germany, Ireland, Great Britain, Sweden, Denmark, Italy, Finland and in some countries in capital regions only, or these including the country capital, e.g. in Portugal, the Czech Republic, Spain and Hungary (see Figure 2).

CONCLUSIONS

The article suggests a new analytical approach in the assessment of the areas and directions of changes in smart specialization development referring to the European regions. It assumes the application of fuzzy classification methods and a dynamic approach (regions as so-called taxonomic units). It allowed for the single identification of four classes of regions in the entire adopted research period and therefore to avoid the problem of selecting an optimal number of classes for each of the eleven years. It also served as the starting point in the analysis of the level of regions' membership in particular classes in each year covered by the adopted research period.

The analysis results facilitated the identification of regions characterized by the changing sector system and maintaining the sector specialization profile (almost 60% of all EU regions). The class featuring agricultural specialization covered 13 regions, a service specialization class – 72 regions, a multifunctional class – 35 and an industrial specialization class – 36 regions. Therefore the obtained results allowed for the positive verification of the thesis put forward in the introduction of the article about the usefulness of the suggested approach in the identification of areas and changes of smart specialization development in the European regions.

Another step towards smart specialization identification should focus on narrowing down these sectors and sections in which regions show smart specialization.

An interesting and possible development of the obtained classification results may be assigning regions to classes with smart specializations, which have already been identified in many EU regions, as well as an observation as to what extent the high employment share in a particular sector – especially KIS 2 and HIT 2 – influenced smart specializations determined for particular regions.

REFERENCES

A Strategy for Smart, Sustainable and Inclusive Growth, Communication from the Commission, Europe 2020, European Commission, COM(2010) 2020 final, Brussels, 2010.

Ahner, D., *Opening Address – Part I*, prepared for Conference "Regions for Economic Change. Fostering Smart and Sustainable Growth in Cities and Regions", Brussels, pp. 23–24 June, 2011.

Aiginger, K, Pfaffermayr, M., *The Single Market and Geographic Concentration in Europe*, "Review of International Economics", 12(1), pp. 1–11, 2004.

Aiginger, K., Davies, S. W., *Industrial Specialisation and Geographic Concentration: Two Sides of the Same Coin? Not for the European Union*, "Journal of Applied Economics", VII(2), pp. 231–248, 2004.

Alcorta, L., Peres, W., *Innovation Systems and Technological Specialization in Latin America and the Caribbean*, "Research Policy", 26, 1998.

Aldenderfer, M. S., Blashfield, R. K., *Cluster Analysis*. Sage Publications, Newbury Park, CA 1984.

Amiti, M., *Specialisation Patterns in Europe*. "Weltwirtschaftliches Archiv", pp. 573–593, 1999.

Annex IV of the general SF draft regulation, COM (2011) 615.

Bezdek, J. C., Ehrlich, R., Full, W., *FCM*: *The Fuzzy c-Means Clustering Algorithm*, "Computers & Geosciences", pp. 191–203, Vol. 10, no. 2-3, 191-203, 1984.

Bezdek, J. C., *Numerical Taxonomy with Fuzzy Sets*, "Journal of Mathematical Biology", Vol. 1, pp. 57–71, 1974.

Bezdek, J. C., *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York 1981.

Bray, J. R., Curtis, J. T., *An Ordination of the Upland Forest Communities of Southern Wisconsin*, "Ecological Monographies" 27, pp. 325–349, 1957.

Brülhart, M., *Evolving Geographical Concentration of European Union*, "Weltwirtschaftliches Archiv", 137(2), pp. 215–243, 2001.

Brülhart, M., Torstensson, J., *Regional Integration, Scale Economies and Industry Location in the European Union*. CEPR Discussion Paper N. 1435,1996.

Brülhart, M., *Trading Places: Industrial Specialisation in the European Union*, "Journal of Common Market Studies", 36(3), pp. 329–346, 1998.

Comparative Advantage through "Smart" Knowledge-based Specialization: Implications for Science, Council Conclusions on Innovation Union for Europe, 3049th Competitiveness Council meeting. Brussels, 26 November 2010.

Council Conclusions on Innovation for Europe, 3049th Competitiveness Council meeting. Brussels, 26 November 2010.

Cowell, F. A., *Measuring Inequality*. Prentice Hall, London 1995.

Cutrini, E., *The Balassa Index Meets the Dissimilarity Theil Index: A Decomposition Methodology for Location Studies,* "Economic Growth – Institutional and Social Dynamic", Working Papers Series, no. 5, June, 2006.

Cziráky, D., Sambt, J., Rovan, J., Puljiz, J., *Regional Development Assessment: A Structural Equation Approach*, "European Journal of Operational Research", Vol. 174, pp. 427–442, 2005.

David, P., Foray, D., Hall, B., *Measuring Smart Specialisation: The Concept and the Need for Indicators,* 2009, http://cemi.epfl.ch/files/content/sites/cemi/files/users/178044/public/Measuring%20smart%20specialisation.doc.

del Campo, C., Monteiro, C. M. F., Soares, J. O., *The Socioeconomic Diversity of European Regions*, "CES Working Paper Series", no. 131, pp. 1–17, 2006.

Dobrzycka, M., *Inteligentna specjalizacja regionów* [*Intelligent Specialization of Regions*], "Wspólnota" no. 1 of 5th January 2012.

Doświadczenia i szanse regionów [*Regional experiences and chances*], The Ministry for Regional Development, "Przegląd regionalny" no. 2, Warsaw 2008.

Dunn, J. C., *A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact, Well Separated Cluster*, "Journal of Cybernetics", Vol. 3, no. 3, pp. 32–57, 1973.

Europe 2020. A Strategy for Smart and Sustainable Development Fostering Social Inclusion, Communication from the European Commission, COM(2010) 2020, final version, Brussels 2010a.

Europe 2020. Flagship Initiative Innovation Union: Transforming Europe for a Post-crisis World, 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), 2010b.

Everit, B. S., Landau, S., Leese, M., Stahl, D., *Cluster Analysis. 5th Edition*. John Wiley & Sons, Chichester 2011.

Florence, P. S., *Report of the Location of Industry*, "Political and Economic Planning", London, 1939.

Florence, P. S., *The Selection of Industries Suitable for Dispersion into Rural Areas*, "Journal of the Royal Statistical Society", Vol. 107, No. 2, pp. 93−116, 1944.

Foray, D., David, P. A., Hall, B., *Smart Specialisation: The Concept* [in:] *Knowledge for Growth: Prospects for Science, Technology and Innovation*, Report, EUR 24047, European Union, 2009.

Foray, D., Goddard, J., Beldarrain, X. G., Landabaso, M., McCann, P., Morgan, K., Nauwelaers, C., Ortega-Argilés, R., *Guide to Research and Innovation Strategies for Smart Specialisation (RIS 3)*, European Union Regional Policy, March 2012.

Foray, D., *Smart Specialisation: From Academic Idea to Political Instrument, the Surprising Destiny of a Concept and the Difficulties Involved in its Implementation*, prepared for Conference "European Integration Process in the New Regional and Global Settings", Warsaw, October 19th and 20th, 2011.

Foray, D., van Ark, B., *Smart Specialisation in a Truly Integrated Research Area is the Key to Attracting More R&D to Europe*. Knowledge Economists Policy Brief n 1 October 2007.

Fukuyama, Y., Sugeno, M., *A New Method of Choosing the Number of Clusters for the Fuzzy C-means Method*, Proceedings of 5th Fuzzy Systems Symposium, pp. 247–250, 1989.

Gertsbakh, I., Yatskiv, I., *Dynamic Classification of Multidimensional Objects*, "Transport and Telecommunication", Vol. 7, no. 1, 21–27, 2006.

Giannitsis, T., Kager, M., *Technology and Specialization: Dilemmas, Options and Risks*? Expert group "Knowledge for Growth", May 2009.

Godlewska-Majkowska, H. (ed.), *Atrakcyjność inwestycyjna regionów Polski a kształtowanie lokalnych i regionalnych specjalizacji gospodarczych* [*Investment Attractiveness of Polish Regions vs. Establishing Local and Regional Economic Specializations*]. Warsaw School of Economics, Warsaw 2009.

Gordon, A. D., *Classification*. 2nd Edition, Chapman & Hall/CRC, Boca Raton 1999.

Haaland, J. I., Kind, H. J., Midelfart-Knarvik, K. H., Torstensson, J., *What Determines the Economic Geography of Europe*? CEPR Discussion Paper no. 2072, 1999.

Hallet, M., *Regional Specialisation and Concentration in the EU*, Economic papers no. 141, The European Commission, Directorate General for Economic and Financial Affairs, 2000.

Harris, T. R., Stoddard, S. W., Bezdek, J. C., *Application of Fuzzy-set Clustering for Regional Typologies*, "Growth and Change", Vol. 24, no. 2, pp. 155–165, 1993.

Hartigan, J. A, *Clustering Algorithms*. John Wiley & Sons, New York-London-Sydney 1975.

Hill, E. W., Brennan, J. F., Wolman, H. L., *What Is a Central City in the United States? Applying a Statistical Technique for Developing Taxonomies*, "Urban Studies", Vol. 35, no. 11, pp. 1935–1969, 1998.

Höppner, F., *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis, and Image Recognition*. John Wiley&Sons, Chichester 1999.

Jajuga, K., *Statystyczna teoria rozpoznawania obrazów* [*Statistical Theory of Pattern Recognition*]. PWN, Warsaw 1990.

Kaufman, L., Rousseeuw, P. J., *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons, New York 1990.

Ketels, C., Sölvell, O., *Clusters in the EU-10 New Member Countries*, Europe Innova Cluster Mapping, 2006.

Kim, S., *Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in US Regional Manufacturing Structure, 1860-1987*, "The Quarterly Journal of Economics", pp. 881–908, 1995.

Krugman, P., *Geography and Trade*. MIT Press, Cambridge 1991.

Larosse, J., Pontkakis, D., *R&D Specialization: Strategic Intelligence in Priority Setting*, OECD TIP-RHIR Workshop, September 2008.

Lasek, M., *Data Mining. Zastosowania w analizach i ocenach klientów bankowych* [*Data Mining. Applications in Analysis and Assessments of Bank Clients*]. Oficyna Wydawnicza „Zarządzanie i finanse", Biblioteka Menadżera i Bankowca, Warsaw 2002.

Malerba, F., Montobbio, F., *Sectoral Systems and International Technological and Trade Specialisation*, DRUID Summer Conference, Rebild, Denmark, 15–17 June 2000.

Malerba, F., *Sectoral Systems of Innovation: A Framework for Linking Innovation to the Knowledge Base, Structure and Dynamics of Sectors*, "Economics of Innovation and New Technology", 14, 1–2, pp. 63–82, 2005.

Manca, G., Curtin, K., *Fuzzy Analysis for Modelling Regional Delineation and Development: The Case of the Sardinian Mining Geopark*, "Transactions in GIS", Vol. 16, no. 1, pp. 55–79, 2012.

Markowska, M., Jefmański, B., *Fuzzy Classification of European Regions in the Evaluation of Smart Growth,* "Przegląd Statystyczny", Tom 59, 1, pp. 74–93, Warszawa 2012a.

Markowska, M., Jefmański, B., *Ocena pozycji polskich regionów ze względu na inteligentną specjalizację w europejskiej przestrzeni z wykorzystaniem klasyfikacji rozmytej* [The Assessment of Polish Regions with Regard to Smart Specialization in European Space Applying Fuzzy Classification], Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 244, pp. 102–113, Wrocław 2012b.

Markowska, M., Strahl, D., *European Regional Space Classification Regarding Smart Growth Level*, "Comparative Economic Research. Central and Eastern Europe", 4/2012, pp. 233–247.

Masson, M.-H., Denoeux, T., *Belief Functions and Cluster Ensembles*, Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, pp. 323–334. Springer-Verlag Berlin, Heidelberg 2009.

Mazzocchi, M., Montresor, E., *A Multivariate Statistical Approach to the Analysis of Rural Development*, "Agricultural Economics Review", Vol. 1, no. 2, pp. 31–45, 2000.

McCann, P., Ortega-Argilés, R, *Smart Specialisation, Regional Growth and Applications to EU Cohesion Policy*. Groningen University 2001.

Midelfart, K. H., Overman, H. G., Redding, S. J., Venables, A. J., *The Location of European Industry* [in:] Dierx, A., Ilzkovitz, F., Sekkat, K. (eds.) *European Integration and the Functioning of Product Markets*, pp. 113–170. Edward Elgar 2004.

Mulligan, G. F., Schmidt, C., *A Note on Localisation and Specialization*, "Growth and Change", 36(4), pp. 565–576, 2005.

NACE Rev. 2. Statistical Classification of Economic Activities in the European Community. Methodologies and Working Papers, European Commission, Luxembourg: Office for Official Publications of the European Communities, 2008.

Nowak, E., *Rozmyta analiza skupień jako metoda klasyfikacji podregionów ze względu na aktywność gospodarczą* [*Fuzzy Cluster Analysis as the Method for Sub-regions Classification Regarding Economic Activity*], "Rozwój Regionalny i Polityka Regionalna", no. 3, pp. 21–36, 2008.

Pettersson, Ö., *Microregional Fragmentation in a Swedish County*, "Papers in Regional Science", Vol. 80, pp. 389–409, 2001.

Proposal for a Regulation of the European Parliament and of the Council on support for rural development by the European Agricultural Fund for Rural Development (EAFRD), European Commission, Brussels, 19.10.2011 COM(2011) 627 final/2, 2011/0282 (COD).

Regional Policy Contributing to Smart Growth in Europe 2020, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, SEC, 2010, 1183. COM(2010) 553 final, European Commission, Brussels, 6.10.2010.

Regions in the European Union. Nomenclature of Territorial Units for Statistics NUTS 2006/EU-27, Series: Methodologies and Working Papers, European Commission, Luxemburg 2007.

Rovan, J., Sambt, J., *Socio-economic Differences among Slovenian Municipalities: A Cluster Analysis Approach*, "Developments in Applied Statistics", Vol. 19, pp. 265–278, 2003.

Sapir, A., *The Effects of Europe's Internal Market Program on Production and Trade: A First Assessment*, "Weltwirtschaftliches Archiv", (132), pp. 457–475, 1996.

Soares, O. J., Marquês, M. M. L., Monteiro, C. M. F., *A Multivariate Methodology to Uncover Regional Disparities: A Contribution to Improve European Union and governmental Decisions*, "European Journal of Operational Research", Vol. 145, pp. 121–135, 2003.

Specjalizacja ekonomiczna regionów i jej implikacje dla rozwoju regionalnego [*Economic Specialization of Regions and Its Implications for Regional Growth*], "Przegląd Regionalny", no. 2, September 2008, The Ministry for Regional Development, The Department for Regional Programmes Coordination.

Stimson, R., Baum, S., Mullins, P., O'Connor, K., *A Typology of Community Opportunity and Vulnerability in Metropolitan Australia*, "Papers in Regional Science", Vol. 80, pp. 45–66, 2001.

Strahl, D. (ed.), *Metody oceny rozwoju regionalnego* [*Regional Development Assessment Methods*]. University of Economics Publishing House, Wrocław 2006.

Topaloglou, L., Kallioras, D., Manetos, P., Petrakos, G., *A Border Regions Typology in the Enlarged European Union*, "Journal of Borderlands Studies", Vol. 20, no. 2, pp. 67–89.

van Zeebroeck, N., van Pottelsberghe de la Potterie, B., Han, W., *Issues in Measuring the Degree of Technological Specialisation with Patent Data*, "Scientometrics", Vol. 66, no. 3, pp. 481–492, 2006.

Wang, W., Zhang, Y., *On Fuzzy Cluster Validity Indices*, "Fuzzy Sets and Systems", Vol. 158, issue 19, pp. 2095–2117, 2007.

Wu, K. L., Yang, M. S., *A Cluster Validity Index for Fuzzy Clustering*, "Pattern Recognition Letters", Vol. 26, issue 9, pp. 1275–1291, 2005.

Wysocki, F., *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich* [*Taxonomic Methods for Exploring Economic Types in Farming and Rural Areas*]. University of Life Sciences in Poznań Publishing House, Poznań 2010.

Xie, X. L., Beni, G. A., *Validity Measure for Fuzzy Clustering*, "IEEE Transactions on Pattern Analysis and Machine Intelligence", Vol. 13, no. 8, pp. 841–847, 1991.

Zeliaś, A. (ed.), *Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym* [*Taxonomic Analysis of Spatial Diversification of Living Standard in Poland in a Dynamic Perspective*]. Cracow University of Economics Publishing House, Kraków 2000.

http://ec.europa.eu/agriculture/cap-post-2013/legal-proposals/com627/627_en.pdf
http://ec.europa.eu/eurostat
http://ec.europa.eu/eurostat/statistics-explained/index.php/Employment_statistics
http://s3platform.jrc.ec.europa.eu

Received: May 2012, revised: February 2015

Abbreviations used in the study: EU 27 (regions from all EU countries), EU 12 (regions from countries of the two recent accessions), EU 15 (regions from old-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 (Luxembourg), PT (Portugal), SE (Sweden), UK (United Kingdom), 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).

APPENDIX

Table 1

Basic statistics of the analyzed variables in the period of 2000–2010

Variable	Statistic	Year										
		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
AGRI	max	63.10	61.94	53.76	54.03	51.84	52.51	49.77	49.22	48.48	50.03	50.49
	min	0.32	0.19	0.29	0.27	0.22	0.20	0.49	0.33	0.16	0.30	0.32
	median	5.17	4.87	4.60	4.56	4.54	4.33	4.38	4.05	4.01	3.69	3.83
	mean	8.39	8.08	7.79	7.63	7.19	7.08	6.87	6.53	6.47	6.13	6.12
	standard deviation	9.78	9.46	8.79	8.79	8.06	7.99	7.70	7.31	7.42	7.08	7.15
	coefficient of variation	116.6	117.1	113.0	115.2	112.1	112.8	112.1	112.0	114.6	115.6	116.8
	max/min	197.2	326.0	185.4	200.1	235.6	262.6	101.6	149.2	303.0	166.8	157.8
MANUF	max	45.26	46.19	46.62	45.57	44.76	46.29	46.06	46.19	46.82	44.64	44.04
	min	12.32	10.79	10.64	10.96	10.18	10.92	12.16	10.78	9.38	8.47	7.19
	median	28.21	28.29	28.18	27.46	26.96	26.38	26.28	26.51	25.64	24.49	24.13
	mean	28.83	28.49	28.24	27.67	27.45	27.10	27.04	27.29	26.83	25.60	25.18
	standard deviation	6.97	6.81	6.88	6.78	6.77	6.65	6.78	6.98	7.09	7.07	7.13
	coefficient of variation	24.2	23.9	24.3	24.5	24.7	24.5	25.1	25.6	26.4	27.6	28.3
	max/min	3.7	4.3	4.4	4.2	4.4	4.2	3.8	4.3	5.0	5.3	6.1
SERV	max	87.34	89.02	89.07	88.68	89.60	88.51	87.30	88.35	90.18	90.49	92.15
	min	19.44	20.72	24.80	24.52	26.70	26.99	29.10	29.37	30.38	29.72	30.01
	median	63.87	64.45	65.20	66.09	66.53	66.73	66.48	67.69	68.68	69.06	69.56
	mean	62.78	63.44	63.97	64.70	65.36	65.82	66.09	66.18	66.70	67.88	68.29
	standard deviation	11.03	10.71	10.84	10.77	10.57	10.37	10.29	10.26	10.59	10.05	9.93
	coefficient of variation	17.56	16.89	16.94	16.64	16.17	15.76	15.57	15.50	15.88	14.81	14.54
	max/min	4.49	4.30	3.59	3.62	3.36	3.28	3.00	3.01	2.97	3.04	3.07
KIS 2	max	74.59	75.47	81.85	80.83	79.95	72.44	73.35	71.61	73.53	69.04	69.48
	min	23.84	24.89	24.85	22.98	26.80	24.30	26.55	26.62	25.20	30.28	36.47
	median	45.52	45.30	45.97	46.70	47.00	47.41	47.33	48.24	48.07	54.53	54.85
	mean	45.90	46.23	46.46	46.92	47.41	47.45	47.85	48.01	47.97	54.39	54.71
	standard deviation	8.06	8.26	8.21	8.25	7.76	7.82	7.83	8.00	7.62	6.77	6.44
	coefficient of variation	17.57	17.86	17.67	17.58	16.37	16.47	16.37	16.67	15.89	12.45	11.78
	max/min	3.13	3.03	3.29	3.52	2.98	2.98	2.76	2.69	2.92	2.28	1.90
HIT 2	max	56.69	50.14	51.22	52.84	53.37	55.40	52.65	57.61	58.97	51.25	53.06
	min	2.34	2.91	2.95	2.12	3.50	3.20	3.49	3.02	3.87	3.25	2.37
	median	22.78	23.27	23.03	22.46	22.85	22.90	22.03	22.50	22.48	19.09	19.39
	mean	23.05	23.29	23.01	22.71	22.80	22.59	22.71	22.70	22.93	20.35	20.38
	standard deviation	9.83	9.86	9.64	9.54	9.43	9.32	9.88	9.58	9.97	9.71	9.57
	coefficient of variation	42.65	42.31	41.90	42.03	41.36	41.27	43.52	42.19	43.46	47.71	46.97
	max/min	24.23	17.22	17.35	24.89	15.23	17.31	15.07	19.09	15.22	15.78	22.34

Source: authors' calculations based on Eurostat data: http://ec.europa.eu/eurostat (downloaded on January 15, 2012)

Table 2

Values of indices for fuzzy classification quality

Measure	Number of classes				
	2	**3**	**4**	**5**	**6**
Xie-Beni Index	0.000165	0.000148	**0.000146**	0.000206	0.000156
Fukuyama-Sugeno Index	-3678.26	-5805.4	**-7516.44**	-7003.72	-7141.63
Partition coefficient	0.624208	**0.64268**	0.635074	0.585148	0.57115
Partition entropy	**0.000151**	0.000134	**0.000128**	0.000136	0.000137

Source: authors' calculations

Table 3

Arithmetic means in classes obtained using c-means fuzzy classification method

Class no.	Variable				
	AGRI	***MANUF***	***SERV***	***KIS 2***	***HIT 2***
1	30.37	25.25	44.39	39.46	13.09
2	3.05	22.16	74.59	55.43	24.43
3	9.05	27.78	63.15	41.77	14.36
4	4.47	34.97	60.61	47.66	30.55

Source: authors' calculations

Table 4

Number of regions from EU countries in classes obtained using c-means fuzzy classification method and for classes obtained at 0.8 membership threshold

Country (number of regions)	**Number of object-periods in a class and membership thresholds**				**Number of object-periods in a class and membership thresholds**				**Undefined**
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**	
	0.332 –0.999	0.377 –0.999	0.335 –0.999	0.343 –0.999	0.800 –0.999	0.800 –0.999	0.800 –0.999	0.800 –0.999	< 0.800
Austria (9)		18	53	28		13	34	14	38
Belgium (11)		98	4	19		73		6	42
Germany (39)		117	45	267		64	14	196	155
Denmark (5)		55				37			18
Spain (17)	1	12	132	42		10	81	35	61
Finland (5)		46		9		23		1	31
France (22)		115	50	77		70	20	39	113
Greece (13)	76	2	65		46		9		88
Ireland (2)		13	8	1		8	3		11
Italy (21)		29	120	82		14	74	61	82
Luxembourg (1)		10	1						11
The Netherlands (12)		131		1		112			20
Portugal (7)	10	9	58		4	1	22		50
Sweden (8)		86		2		74			14
United Kingdom (37)		362	18	27		305	6	6	90
Bulgaria (6)	4		57	5			27		39
Cyprus (1)			11				9		2
The Czech Republic (8)		11	5	72		11		49	28
Estonia (1)			11				8		3
Hungary (7)		11	19	47		8	8	36	25
Lithuania (1)			11				4		7
Latvia (1)			11				9		2
Malta (1)		4	3	4		2		1	8
Poland (16)	56		104	16	43		24	5	104
Romania (8)	73	2	7	6	56		5		27
Slovenia (2)		2		20		1		9	12
Slovakia (4)		11	3	30		11		7	26
EU 27 (265)	**220**	**1144**	**796**	**755**	**149**	**837**	**357**	**465**	**1107**
EU 15 (209)	**87**	**1103**	**554**	**555**	**50**	**804**	**263**	**358**	**824**
EU 12 (56)	**133**	**41**	**242**	**200**	**99**	**33**	**94**	**107**	**283**

Source: authors' calculations

Table 5

Regions included in the same class at all years of 2000–2010 period

Class	Number of regions (country)	Regions
1	**13** (GR 3), (PL 4), (RO 6)	(GR) Anatoliki Makedonia, Thraki, Dytiki Makedonia, Peloponnisos, (PL) Lubelskie, Podkarpackie, Świętokrzyskie, Podlaskie, (RO) Nord-Vest, Centru, Nord-Est, Sud-Est, Sud–Muntenia, Sud-Vest Oltenia
2	**72** (BE 7), (CZ 1), (DK 5), (DE 4), (IE 1), (ES 1) (FR 3), (IT 2) , (HU 1), (NL 11), (AT 1), (SK 1), (FI 4), (SE 7), (UK 23)	(BE) **Région de Bruxelles-Capitale,** Prov. Vlaams-Brabant, Prov. Brabant Wallon, Prov. Hainaut, Prov. Liège, Prov. Luxembourg, Prov. Namur, (CZ) **Praha**, (DK) **Hovedstaden**, Sjælland, Syddanmark, Midtjylland, Nordjylland, (DE) **Berlin,** Bremen, Hamburg, Schleswig-Holstein, (IE) **Southern and Eastern**, (ES) **Comunidad de Madrid**, (FR) **Île de France**, Midi-Pyrénées, Provence-Alpes-Côte d'Azur, (IT) Liguria, Lazio, (HU) **Közép-Magyarország**, (NL) Groningen, Friesland, Drenthe, Overijssel, Gelderland, Flevoland, Utrecht, **Noord-Holland,** Zuid-Holland, Zeeland, Noord-Brabant, (AT) **Wien**, (SK) **Bratislavský kraj**, (FI) Itä-Suomi, **Etelä-Suomi,** Pohjois-Suomi, Åland, (SE) **Stockholm,** Sydsverige, Östra Mellansverige, Västsverige, Norra Mellansverige, Mellersta Norrland, Övre Norrland, (UK) Northumberland and Tyne and Wear, Greater Manchester, Merseyside, North Yorkshire, South Yorkshire, West Yorkshire, East Anglia, Bedfordshire and Hertfordshire, Essex, Berkshire, **Inner London, Outer London,** Buckinghamshire and Oxfordshire, Surrey, East and West Sussex, Hampshire and Isle of Wight, Kent, Gloucestershire, Wiltshire and Bristol, Dorset and Somerset, Devon, West Wales and The Valleys, East Wales, Eastern Scotland, South Western Scotland, North Eastern Scotland
3	**35** (BG 3), (EE 1), (GR 2), (ES 10), (IT 7), (CY), (LV), (LT), (HU 1), (AT 1), PL 3), (PT 4)	(BG) Severozapaden, Yugoiztochen, **Yugozapaden**, (EE) Eesti, (GR) Kentriki Makedonia, Notio Aigaio, (ES) Principado de Asturias, Cantabria, Castilla y León, Castilla-la Mancha, Extremadura, Comunidad Valenciana, Illes Balears, Andalucía, Región de Murcia, Canarias, (IT) Provincia Autonoma Bolzano, Toscana, Campania, Puglia, Calabria, Sicilia, Sardegna, (CY) Cypr, (LV) Latvija, (LT) Lietuva, (HU) Dél-Alföld, (AT) Tirol, (PL) **Mazowieckie**, Zachodniopomorskie, Warmińsko-Mazurskie, (PT) Norte, Algarve, Alentejo, Região Autónoma dos Açores
4	**36** (CZ 5), (DE 14) (ES 3), (FR 1) (IT 6), (HU 3), (AT 2), (SI 1) (SK 1)	(CZ) Stredni Cechy, Jihozápad, Severovýchod, Jihovýchod, Stredni Morava, (DE) Stuttgart, Karlsruhe, Freiburg, Tübingen, Niederbayern, Oberpfalz, Oberfranken, Mittelfranken, Unterfranken, Braunschweig, Schwaben, Detmold, Arnsberg, Chemnitz, (ES) País Vasco, Comunidad Foral de Navarra, Cataluña, (FR) Franche-Comté, (IT) Piemonte, Lombardia, Veneto, Friuli-Venezia Giulia, Emilia-Romagna, Marche, (HU) Közép-Dunántúl, Nyugat-Dunántúl, Észak-Magyarország (AT) Vorarlberg, Oberösterreich, (SI) Vzhodna Slovenija, (SK) Západné Slovensko

Country capitals or regions including country capital are in bold.

Source: authors' calculations

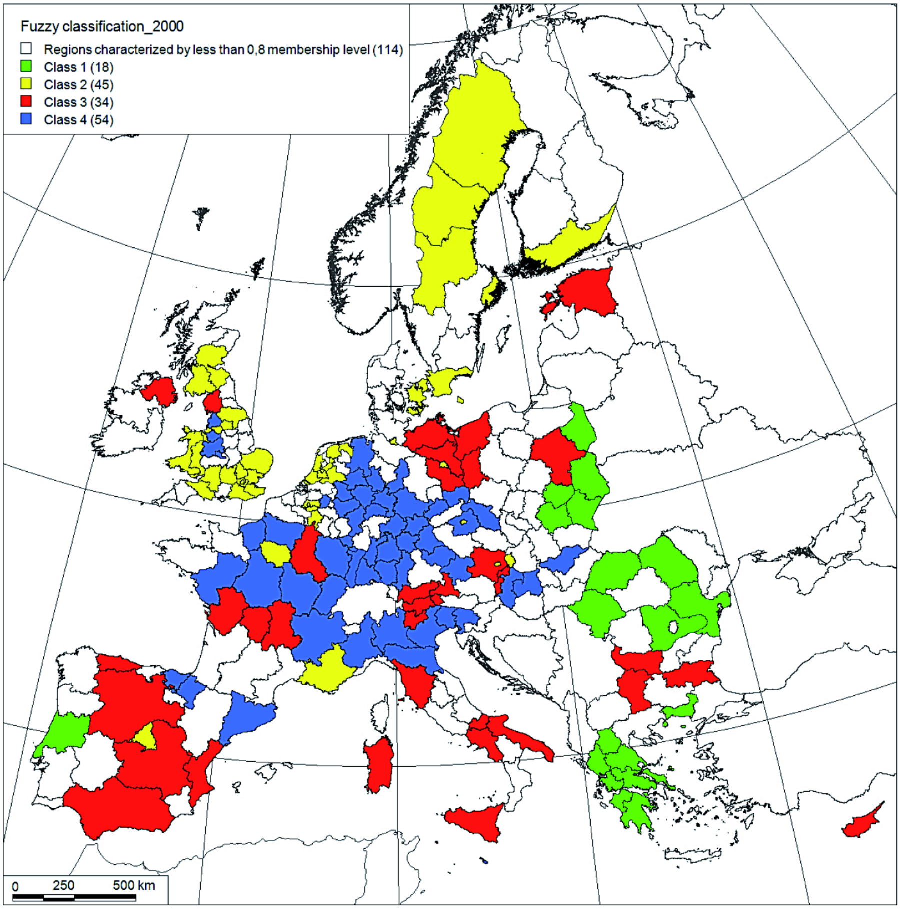

Figure 1. The results of European regions classification (2000)

Source: Graph by MapInfo Professional 10.5, based on own study results

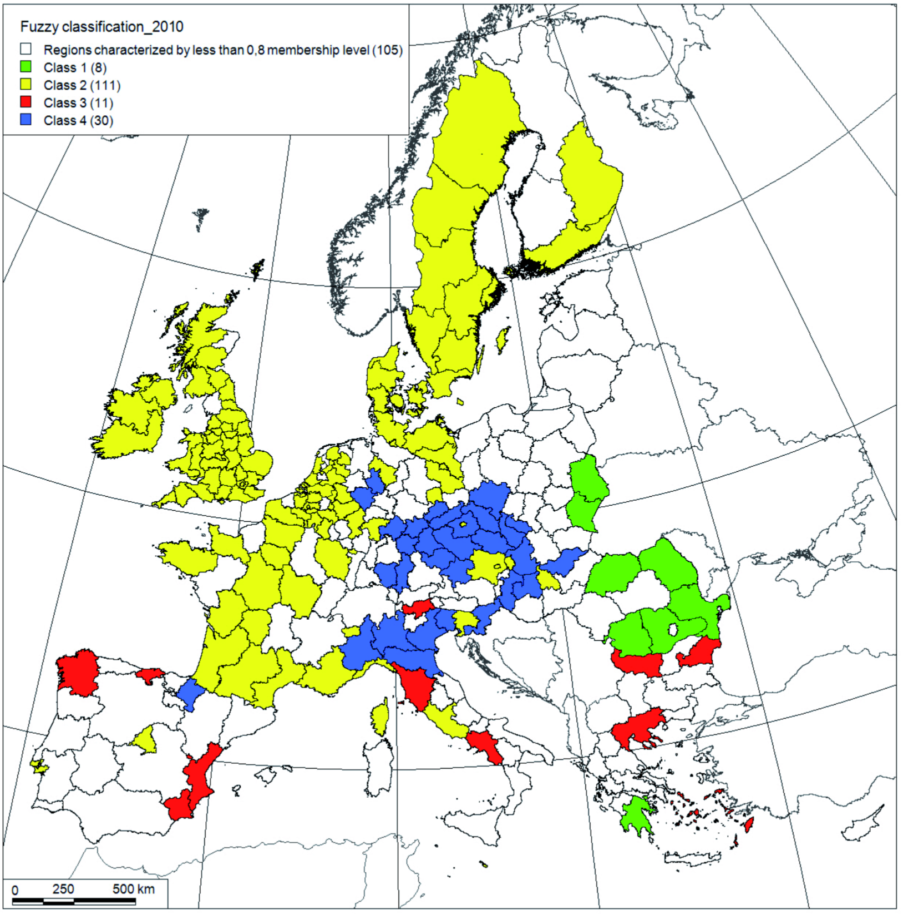

Figure 2. The results of European regions classification (2010)

Source: Graph by MapInfo Professional 10.5, based on own study results