

Alicja KOLASA-WIĘCEK

Politechnika Opolska, Wydział Ekonomii i Zarządzania, Katedra Ekonomii i Badań Regionalnych
ul. Waryńskiego 4; 45-047 Opole; e-mail: a.kolasa-wiecek@po.opole.pl

Marek TUKIENDORF

Politechnika Opolska, Wydział Mechaniczny, Katedra Techniki Rolniczej i Leśnej
ul. Mikołajczyka 5; 45-271 Opole; e-mail: m.tukiendorf@po.opole.pl

APPLICATION OF AGGLOMERATIVE METHOD OF CLUSTERING IN SEGMENTATION OF EUROPEAN UNION COUNTRIES

Summary

Research aimed to grouping the EU 27 states in homogeneous sets as regards similarities in the capacities and capabilities of crop agricultural production and ranking of Polish position in the context of this issue. The agglomerative method of clustering was used for this purpose. The states with high rates of harvesting of yields per hectare, i.e. Austria, Belgium, Holland, Germany, France, Sweden, Slovenia, Ireland, Finland, Great Britain, Italy constitute a numerous group admitted to one cluster. Poland has a high proportion of arable land in EU and also one of the lowest rate of the yields per hectare. The study was conducted using the Statistica 10.0.

Key words: farming; arable land; crop production; crop yield; analysis of concentrations; agglomerative method; statistical surveys; European Union

ZASTOSOWANIE METOD GRUPOWANIA AGLOMERACYJNEGO W SEGMENTACJI PAŃSTW UNII EUROPEJSKIEJ

Streszczenie

Badania miały na celu pogrupowanie państw UE 27 w jednorodne zbiory pod względem podobieństw w obszarze zdolności i możliwości roślinnej produkcji rolnej oraz uszeregowanie pozycji Polski w kontekście tego zagadnienia. Wykorzystano w tym celu metody grupowania aglomeracyjnego. Liczną grupę zakwalifikowaną w jedno skupienie stanowią państwa z wysokim wskaźnikiem pozyskiwania plonów z hektara upraw, tj. Austria, Belgia, Holandia, Niemcy, Francja, Szwecja, Słowenia, Irlandia, Finlandia, Wielka Brytania, Włochy. Polska charakteryzuje się wysokim udziałem gruntów ornych i jednocześnie posiada jedną z najniższych w Unii wartość uzyskiwania plonu z hektara. Badania przeprowadzono z wykorzystaniem pakietu Statistica 10.0.

Słowa kluczowe: rolnictwo; grunty orne; produkcja roślinna; plony roślin; analiza skupień; metoda aglomeracyjna; badania statystyczne; Unia Europejska

1. Wprowadzenie

Zboża są podstawowym surowcem rolnym. Odgrywają zasadniczą rolę stanowiąc najważniejszy produkt spożywczy i decydując o bezpieczeństwie żywnościowym danego kraju [10]. W krajach Unii Europejskiej zboża stanowią najważniejszą grupą roślin uprawnych. [3, 6]. Powierzchnia zasiewu zbóż w krajach członkowskich UE 27 w 2009 roku stanowiła 50,1%. [9]. Polska ma w tym kontekście szczególny udział, ponieważ powierzchniowo zajmuje drugą po Francji (9393 tys. ha) pozycję z 8583 tys. ha, a trzecią w kolejności stanowią Niemcy z arealem 6908 tys. ha. Pod względem zbiorów upraw w 2009 roku najwyższy udział przypadł Francji – 23,5%, Niemcom – 16,7% oraz Polsce 10% [9]. Kraje te odpowiadają za ponad 50% zbiorów w krajach Unii. Od lat Polska zajmuje czołowe pozycje w produkcji m.in. takich artykułów rolnych jak żyto, owies, buraki cukrowe, rzepak, pszenica, czy ziemniaki.

Podjęte badania miały na celu pogrupowanie państw UE 27 w jednorodne zbiory pod względem podobieństw w obszarze zdolności i możliwości roślinnej produkcji rolnej oraz uszeregowanie pozycji Polski w kontekście tego zagadnienia.

2. Metodyka badań

Analiza skupień ma za zadanie połączenie zbioru obiektów w homogeniczne grupy. Grupowania obiektów dokonano w oparciu o pięć metod hierarchicznych aglomeracyj-

nych: metodę pojedynczego wiązania, wiązania pełnego, metodę średnich połączeń, średnich połączeń ważonych i metodę Warda. Zasada działania wg procedury aglomeracyjnej polega na tworzeniu macierzy podobieństw klasyfikowanych obiektów, a w kolejnych krokach na łączeniu w skupienia obiektów najbardziej do siebie podobnych. W początkowym etapie grupowania przyjmuje się, iż każdy obiekt stanowi odrębne skupienie. W kolejnym etapie procedury osłabione zostaje kryterium uznawania obiektów za takie same, co w konsekwencji prowadzi do grupowania elementów w grupy coraz bardziej różniące się od siebie do momentu, kiedy wszystkie zostaną połączone.

W procesie klasyfikowania obiektów w grupy zastosowano pięć metod wiązania [2, 5]:

- metoda pojedynczego wiązania, odległość między dwoma skupieniami jest zdefiniowana jako odległość między dwoma najbliższymi punktami, po jednym z każdego skupienia,
- metoda pełnego wiązania, odległość między dwoma skupieniami jest zdefiniowana jako odległość między dwoma najbardziej oddalonymi punktami, po jednym z każdego skupienia,
- metoda średnich połączeń (średniej grupowej), w której odległość między dwiema grupami równa jest średniej odległości pomiędzy wszystkimi parami obiektów należących do tych skupień,
- metoda średnich połączeń ważonych, zbliżona do metody średnich połączeń, przy czym w procedurze obliczeniowej przyjmuje się wagi proporcjonalne do liczby obiektów zawartych w danym skupieniu,

– metoda Warda, wykorzystująca w tym celu analizę wariancji.

Analiza skupień dąży do heterogeniczności skupień poprzez łączenie ze sobą obiektów na podstawie wyliczonych odległości między nimi. Spośród wielu możliwych definicji określania odległości przyjęto powszechnie stosowaną odległość euklidesową:

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{i,k} - x_{j,k})^2} \quad (1)$$

gdzie:

d_{ij} – odległość pomiędzy i-tym a j-tym obiektem,

$x_{i,k}$ – wartość k-tej cechy dla i-tego obiektu,

$x_{j,k}$ – wartość k-tej cechy dla j-tego obiektu,

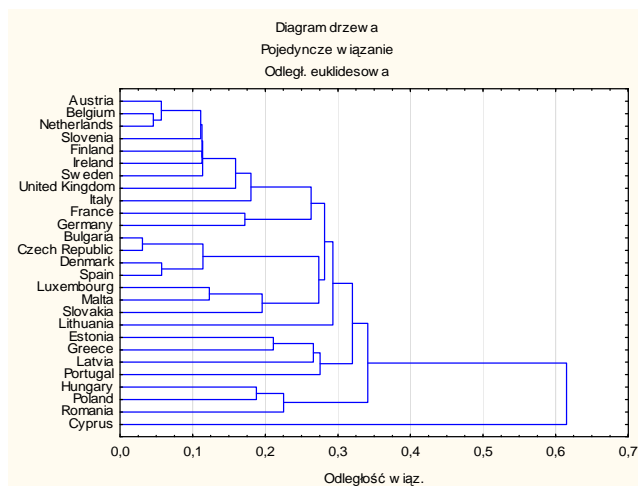
n – liczba cech charakteryzujących obiekty.

Do analizy wykorzystano dane z lat 2009 i 2010 posiłkując się bazą FAO oraz rocznikiem statystycznym [1, 9]. Są to dane obejmujące: udział gruntów ornych w % powierzchni ogólnej kraju, udział zbiorów upraw na tle Unii w %, uzyskane plony z 1 ha oraz udział ludności rolniczej w % ogółu ludności kraju. Analizy prowadzono z wykorzystaniem pakietu Statistica 10.0. Przyjęte do badań dane wyjściowe wyrażone są w różnych jednostkach miary, toteż wstępnie poddano je normalizacji poprzez powszechnie stosowaną metodę standaryzacji danych [7].

3. Wyniki badań i ich interpretacja

Z zastosowanych metod klasyfikacyjnych analizy aglomeracyjnej, identyczne wyniki otrzymano w przypadku metod pojedynczego wiązania, średnich połączeń oraz średnich połączeń ważonych. Kraje UE-27 zostały sklasyfikowane w następujące skupiska:

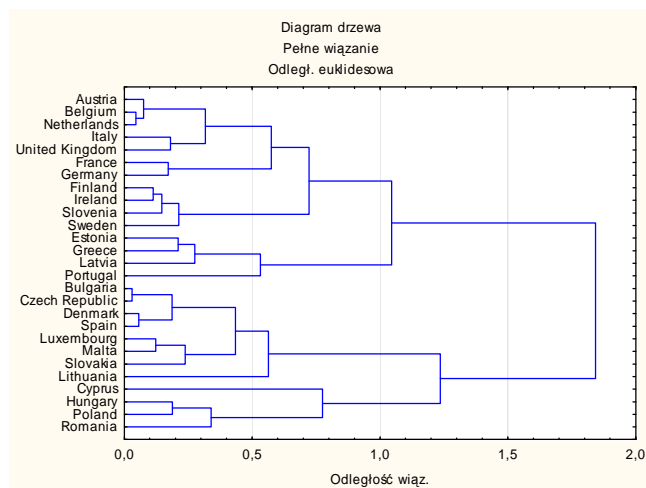
1. Rumunia, Polska, Węgry, Cypr,
2. Portugalia, Łotwa, Grecja, Estonia,
3. Słowacja, Malta, Luksemburg, Litwa, Hiszpania, Dania, Czechy, Bułgaria,
4. Niemcy, Francja, Szwecja, Słowenia, Irlandia, Finlandia, Wielka Brytania, Włochy, Holandia, Belgia, Austria.



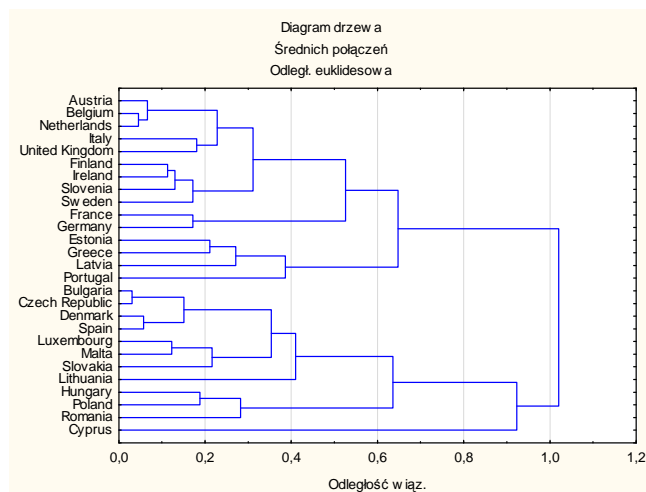
Źródło: opracowanie własne
Rys. 1. Dendrogram uzyskany metodą pojedynczego wiązania
Fig. 1. Dendrogram for the single linkage clustering

Na podstawie dwóch pozostałych metod, tj. Warda oraz pełnego wiązania, kraje pogrupowane zostały dokładnie wg schematu jak powyżej z tą różnicą, że skupienia 2 i 3 wymieniły się grupami krajów.

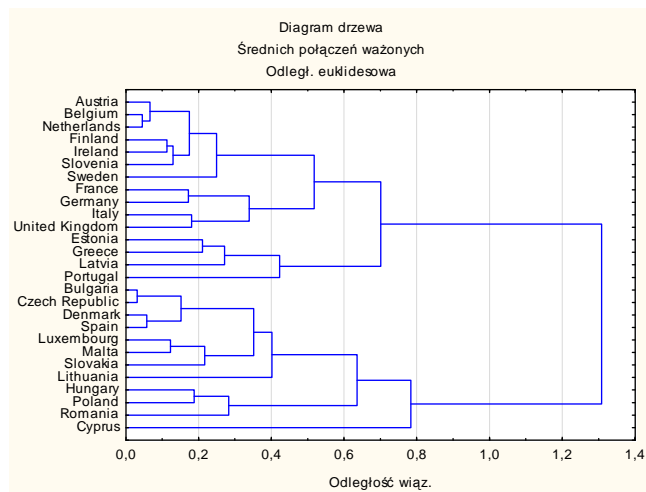
Efektom grupowania metodą aglomeracyjną jest drzewo hierarchiczne tzw. dendrogram. Poniższe rysunki prezentują wyniki klasyfikacji dla poszczególnych metod.



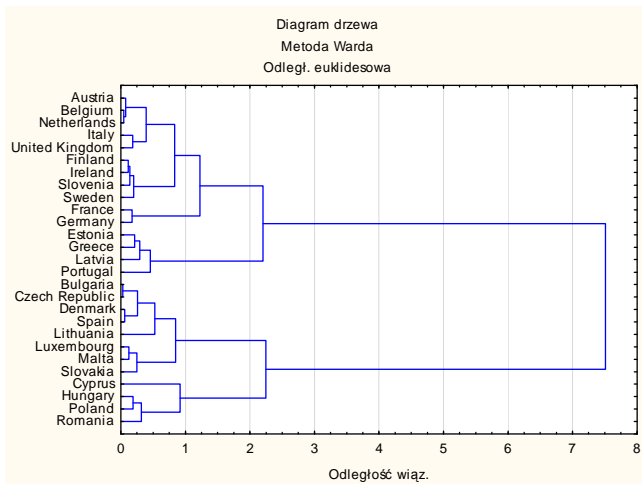
Źródło: opracowanie własne
Rys. 2. Dendrogram uzyskany metodą pełnego wiązania
Fig. 2. Dendrogram for the complete linkage clustering



Źródło: opracowanie własne
Rys. 3. Dendrogram uzyskany metodą średnich połączeń
Fig. 3. Dendrogram for average linkage clustering



Źródło: opracowanie własne
Rys. 4. Dendrogram uzyskany metodą średnich połączeń ważonych
Fig. 4. Dendrogram for weighted-average linkage clustering



Źródło: opracowanie własne

Rys. 5. Dendrogram uzyskany metodą Warda

Fig. 5. Dendrogram for Ward's method

Wyniki klasyfikacji zestawiono w tab. 1.

Tab. 1. Wyniki grupowania państw UE metodami aglomeracyjnymi

Table 1. The results of grouping the EU countries by agglomerative method

Lp.	Kraje UE	Metoda grupowania				
		Metoda pojedynczego wiązania	Metoda pełnego wiązania	Metoda Warda	Metoda średnich połączeń	Metoda średnich połączeń ważonych
1	Austria	4	4	4	4	4
2	Belgia	4	4	4	4	4
3	Bułgaria	3	2	2	3	3
4	Cypr	1	1	1	1	1
5	Czechy	3	2	2	3	3
6	Dania	3	2	2	3	3
7	Estonia	2	3	3	2	2
8	Finlandia	4	4	4	4	4
9	Francja	4	4	4	4	4
10	Niemcy	4	4	4	4	4
11	Grecja	2	3	3	2	2
12	Węgry	1	1	1	1	1
13	Irlandia	4	4	4	4	4
14	Włochy	4	4	4	4	4
15	Łotwa	2	3	3	2	2
16	Litwa	3	2	2	3	3
17	Luxemburg	3	2	2	3	3
18	Malta	3	2	2	3	3
19	Holandia	4	4	4	4	4
20	Polska	1	1	1	1	1
21	Portugalia	2	3	3	2	2
22	Rumunia	1	1	1	1	1
23	Słowacja	3	2	2	3	3
24	Słowenia	4	4	4	4	4
25	Hiszpania	3	2	2	3	3
26	Szwecja	4	4	4	4	4
27	W. Brytania	4	4	4	4	4

Źródło: opracowanie własne

Analizując dendrogramy poszczególnych metod zauważa się homogeniczność jednego z krajów. Na tle skupiska 1, państw takich jak Polska, Rumunia, Węgry, Cypr, ten ostatni stanowi punkt odstający. Powodem uzyskanej klasyfikacji jest najniższa wśród państw członkowskich UE, ale także i rozpatrywanej grupy, wartość uzyskanych plonów

z 1 ha. Na Cyprze wynosi ona 9 dt/ha, przy średniej w Unii odnotowanej na poziomie 48 dt/ha [9]. Skupienie pozostałych krajów tej grupy charakteryzuje podobieństwo w obszarze wielkości udziału zbiorów na tle UE (6,5% wobec średniej unijnej wynoszącej 3,7%). Ponadto skupienie charakteryzuje stosunkowo wysoki udział gruntów ornych, na poziomie odpowiednio w Polsce – 38,7%, w Rumuni – 36,9% oraz na Węgrzech – 49,5 i jest wyższy od średniej europejskiej kształtującej się w granicach 25,7%.

W skupieniu 2 znalazły się państwa sklasyfikowane względem podobieństwa w zakresie wysokiego udziału ludności rolniczej w kraju, tj. w przedziale 9,1-10,6% ogółu ludności, wobec średniej europejskiej 5,4%. Skupienie charakteryzuje również bardzo niski udział w zbiorach zbóż danego kraju w odniesieniu do średniej unijnej, o wartościach od 0,3% uzyskanej w Estonii do 1,6% w Grecji, przy średniej europejskiej na poziomie 3,7%.

Kraje zgrupowane w skupieniu 3 charakteryzuje wysoki udział gruntów ornych o średniej wartości na poziomie ok. 34% przy średniej unijnej wynoszącej 25,7% powierzchni.

Skupienie 4 stanowi najliczniejszą grupę. Obejmując dominującą większość państw wysoko uprzemysłowionych z wysokim udziałem plonów uzyskiwanych z 1 hektara. Średnia wartość uzyskana dla tych państw wynosi niemal 66 dt/ha wobec średniej unijnej na poziomie 50,1 dt/ha. Ponadto państwa tej grupy charakteryzują się niskim udziałem w społeczności ludności rolniczej – 2,7%, na tle średniej europejskiej kształtującej się na poziomie 5,4%.

4. Wnioski

Usystematyzowanie państw względem wybranych cech przy wykorzystaniu aglomeracyjnej analizy skupień umożliwiło rozpoznanie podobieństw między obiektami. Wszystkie z podjętych metod uszeregowały państwa UE w te same grupy państw, z pewnymi odstępstwami w zakresie kolejności wytypowanych skupień. Najliczniejsze skupienie stanowiła grupa państw wysoko uprzemysłowionych, tj. Austria, Belgia, Holandia, Niemcy, Francja, Szwecja, Słowenia, Irlandia, Finlandia, Wielka Brytania, Włochy, które charakteryzuje wysoki wskaźnik uzyskiwania plonów z hektara upraw. Polska, mimo stosunkowo wysokiego udziału arealu gruntów ornych w Europie, nadal charakteryzuje się jedną z najniższych w Unii wartości pozyskiwania plonu z hektara.

5. Bibliografia

- [1] Food and Agricultural Organization <http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor> [dostęp 2.04.2011]
- [2] Gatnar E.: Symboliczne metody klasyfikacji danych. Warszawa: PWN, 1998.
- [3] Ginter A., Szarek S.: Sytuacja dochodowa producentów zbóż na przykładzie uprawy pszenicy, Journal of Agribusiness and Rural Development, 2010, 4(18): 29-39.
- [4] Just M.: Rolnictwo i gospodarka żywnościowa w Polsce. Ministerstwo Rolnictwa i Rozwoju Wsi, Warszawa, 2008.
- [5] Nowak E.: Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych. Warszawa: PWE, 1990.
- [6] Olsen O.: A regional picture of farming in Europe - what, where and how much? Agriculture and fisheries, Statistics in fokus, Eurostat, 44, 2010.
- [7] Ostasiewicz W.: Statystyczne metody analizy danych. Akademia Ekonomiczna we Wrocławiu, 1998.
- [8] Panek T.: Statystyczne metody wielowymiarowej analizy porównawczej. Oficyna Wydawnicza SGH, 2009.
- [9] Rocznik Statystyczny Rolnictwa 2011, Warszawa: GUS, 2011.
- [10] Urban S.: Zmiany na polskim rynku zbóż i ich przyczyny. Roczn. Nauk. 2007, 9, 4: 208-211.