

OCENA ZASTOSOWANIA WYBRANYCH METOD TAKSONOMICZNYCH DO KLASYFIKACJI ZJAWISK SPOŁECZNO-GOSPODARCZYCH

Barbara Prus, Karol Król

Uniwersytet Rolniczy w Krakowie

Streszczenie. Metody taksonomiczne służą do opisu i klasyfikacji złożonych zjawisk społeczno-gospodarczych wyrażonych przy pomocy licznego zestawu zmiennych. Metody te dają różne wyniki, pomimo przyjętego jednakowego zestawu wyjściowych cech badawczych. Celem pracy było przedstawienie wyników klasyfikacji wielocechowej powiatów województwa małopolskiego przy użyciu trzech metod taksonomicznych oraz dokonanie oceny przeprowadzonego podziału. Do badań przyjęto zestaw zmiennych definiujących warunki rozwoju społeczno-gospodarczego. Ocenie poddano trzy metody taksonomiczne – taksonomii przestrzennej, aglomeracyjną metodę Warda oraz metodę Czekanowskiego. Do oceny wykorzystano mierniki oceny klasyfikacji jednostek: indywidualne, homogeniczności oraz heterogeniczności, obliczone dla wydzielonych w wyniku analizy skupień, a także mierniki poprawności skupień. Obliczenia wykonano w programie Taksonomia Numeryczna. W pracy posłużono się metodą opisowo-porównawczą oraz analizą logiczną. Najlepsze wskaźniki homogeniczności świadczące o poprawności wykonania klasyfikacji otrzymano w przypadku metody Warda. Warunek heterogeniczności został najlepiej spełniony w przypadku metody taksonomii przestrzennej. Najlepsze mierniki klasyfikacji, zarówno te świadczące o homogeniczności, jak i heterogeniczności, otrzymano po zastosowaniu metody Czekanowskiego. Przeprowadzone badania potwierdziły najlepszą skuteczność w przypadku metody Warda.

Słowa kluczowe: taksonomia przestrzenna, metoda Warda, metoda Czekanowskiego, rozwój społeczno-gospodarczy, odległość euklidesowa

WSTĘP

Rozwój społeczno-gospodarczy uwarunkowany jest w krajach rozwiniętych poprzez respektowanie zasad rozwoju zrównoważonego, ochrony środowiska [Kauffman i Martin 2017] oraz ładu przestrzennego [Kates i in. 2005]. Definicja rozwoju zrównoważonego

Adres do korespondencji – Corresponding authors: dr inż. Barbara Prus, dr inż. Karol Król, Katedra Gospodarki Przestrzennej i Architektury Krajobrazu, Wydział Inżynierii Środowiska i Geodezji, Uniwersytet Rolniczy im. H. Kołłątaja w Krakowie, ul. Balicka 253c, 30-149 Kraków; e-mail: b.prus@ur.krakow.pl, k.krol@onet.com.pl.

zaistniała po raz pierwszy w 1987 r. w Raportcie Komisji Brundtlanda „Nasza Wspólna Przyszłość” [Our Common Future 1987]. W Polsce Ustawa o ochronie środowiska z 2001 r. definiuje rozwój zrównoważony jako rozwój społeczno gospodarczy, w którym następuje proces integrowania działań politycznych, gospodarczych i społecznych, z zachowaniem równowagi przyrodniczej oraz trwałości podstawowych procesów przyrodniczych [Gwiazdzińska-Goraj i Goraj 2013], w celu zagwarantowania możliwości zaspokajania podstawowych potrzeb poszczególnych społeczności współczesnego i przyszłych pokoleń [Ustawa... 2001]. W celu dostosowania do światowych standardów powstają modele, teorie i analizy połączeń funkcjonalnych w zakresie rozwoju zrównoważonego [Cieszewska 2008]. W 2002 r. na Światowym Szczycie Rozwoju Zrównoważonego przyjęto rozszerzoną definicję pojęcia rozwoju zrównoważonego jako rozwoju, który wymaga równowagi pomiędzy trzema aspektami: społecznym, gospodarczym i środowiskowym [Redclift 2005, Bolis i in. 2017]. Badając uwarunkowania rozwoju należy również wziąć pod uwagę ład przestrzenny, czyli takie ukształtowanie przestrzeni, które tworzy harmonijną całość oraz uwzględnia w uporządkowanych relacjach uwarunkowania i wymagania funkcjonalne, społeczno-gospodarcze, środowiskowe, kulturowe oraz kompozycyjno-estetyczne [Ustawa... 2003]. Zjawiska społeczno-gospodarcze definiowane są w wielowymiarowej przestrzeni zmiennych. Z merytorycznego punktu widzenia opisują zarówno warunki ekonomiczne, struktury i mechanizmy funkcjonowania gospodarki, jak i stan środowiska przyrodniczego [Bagdziński 1994, Gorzelak 1992, Zimny 2007]. Rozwój społeczno-gospodarczy określany jest jako proces pozytywnych zmian ilościowych oraz jakościowych w sferach społecznych, gospodarczych, a także produkcyjno-społecznych, przy pomocy czynników lokalizacyjnych, społeczno-ekonomicznych, a także techniczno-organizacyjnych [Heldak i Raszka 2013, Semenova i in. 2016, Ziemiańczyk 2010]. Z ekonomicznego punktu widzenia rozwój utożsamiany jest z rosnącą liczbą inwestycji, wzrostem produkcji, zatrudnienia i poziomu życia mieszkańców, a także wzrostem poziomu bezpieczeństwa socjalnego i publicznego [Kuciński 2009, Król 2015]. Często rozwój społeczno-gospodarczy jest warunkowany położeniem geograficznym [Szymła 2000]. Natomiast do głównych czynników mających nań wpływ należą działania jednostek samorządu terytorialnego o charakterze administracyjnym, mające na celu m.in. likwidowanie dysproporcji występujących pomiędzy regionami [Siekierski 2005], sterowanie nakładami finansowymi na różne gałęzie gospodarki [Wilkin 2005, Sokołowski 1999], zwiększanie potencjału infrastruktury technicznej, czy np. wybór obszarów do realizacji zadań inwestycyjnych. Podejmowanie działań o charakterze strategicznym wymaga jednak wielowymiarowych analiz czynników mających wpływ na rozwój społeczno-gospodarczy, w tym, w pierwszej kolejności ich odpowiedni, poparty przesłankami merytorycznymi wybór. Cechy opisujące rozwój powinny być charakterystyczne, reprezentatywne oraz niepowtarzalne [Hałasiewicz 2010]. Mają być bowiem podstawą do klasyfikacji obiektów przestrzennych na podgrupy regionów sąsiadujących ze sobą w przestrzeni i wewnątrznie jednorodnych [Assunção i in. 2006].

Wielowymiarową, przestrzenną analizę zjawisk umożliwia zastosowanie procedury klasyfikacyjnej, która polega na wydzieleniu w zbiorze obiektów wielocechowych homogenicznych podgrup [Hellwig 1981]. Z założenia przeprowadzanej klasyfikacji, obiekty należące do tego samego skupienia będą podobnymi (homogenicznymi), zaś należące do różnych skupień – heterogenicznymi [Kolenda 2009]. Metody taksonomii numerycznej

umożliwiają tworzenie struktur przestrzennych [Dunn i Everitt 2004], posiadają szeroką gamę zastosowań [Clifford i Stephenson 1975, Szylar i Prus 2015, Gawroński i in. 2014] m.in. w etnografii [Czekanowski 1913], psychologii [Klavans i Boyack 2017], językoznawstwie, zoologii, botanice [Gutiérrez i Pine 2017], archeologii, geografii ekonomicznej [Tavana i in. 2016], rynkach pracy [Zwick i Syed 2017], waloryzacji rolniczej przestrzeni produkcyjnej [Fierich 1957], geografii [Chojnicki i Czyż 1973, Kostrubiec 1982, Grabiński 1992] czy w zaawansowanych metodach lokalizacyjnych [Ahavar i in. 2016]. Grabiński i in. [1989] opisują możliwości zastosowania metod taksonomicznych w badaniach społeczno-gospodarczych i przestrzennych [Frankowski 1991]. Autorzy ci wprowadzili podział metod taksonomicznych na: hierarchiczne i niehierarchiczne, aglomeracyjne i podziałowe, obszarowe oraz optymalizacyjne [Grabiński i in. 1989]. Metody te różnią się algorytmami obliczeniowymi, wszystkie jednak umożliwiają przeprowadzenie klasyfikacji na wielowymiarowym zestawie zmiennych. Podział na metody hierarchiczne i niehierarchiczne należy do najbardziej ogólnych. Metody hierarchiczne pozwalają na wyodrębnienie pełnej hierarchii tzn. skupienia na poziomie wyższym zawierają rozłączne skupienia poziomów niższych tworząc struktury skupień [Kolenda 2009]. W metodach niehierarchicznych nie jest możliwe przedstawienie procesu tworzenia skupień w postaci struktury. W metodach aglomeracyjnych każdy z obiektów badawczych początkowo jest oddzielnym skupieniem, a proces kończy połączenie wszystkich obiektów w jedną grupę. Metody podziałowe traktują zbiór obiektów jako jedną grupę, kolejne podziały doprowadzają w efekcie końcowym do wyodrębnienia liczby grup równiej liczbie obiektów. W metodach obszarowych hiperprzestrzeń jest dzielona na rozłączne przestrzenie, natomiast w metodach optymalizacyjnych, przy zastosowaniu różnych kryteriów, dochodzi do kolejnych przybliżeń podziału zbioru [Grabiński i in. 1989].

Przedmiotem badań była ocena klasyfikacji przeprowadzonej w obrębie powiatów woj. małopolskiego trzema różnymi metodami taksonomicznymi: metodą taksonomii przestrzennej, Czekanowskiego oraz kombinatoryczną, aglomeracyjną metodą Warda. Wybór metod klasyfikacji obiektów wynikał z założenia przedstawienia wyników szerokiego wachlarza dostępnych procedur obliczeniowych, począwszy od historycznego diagramu i macierzy Czekanowskiego, poprzez metody obszarowe (taksonomii przestrzennej) a na metodach aglomeracyjnych kombinatorycznych (Warda) kończąc. Każda z przedstawionych metod różni się procedurą obliczeniową, natomiast wszystkie można zaliczyć do metod hierarchicznych.

MATERIAŁY I METODY

Wyjściowy zestaw zmiennych determinujących w sposób cząstkowy poziom rozwoju społeczno-gospodarczego ustalono w oparciu o przesłanki merytoryczne (uniwersalność, jakość, interpretowalność oraz sposób oddziaływania zmiennych) oraz ich dostępność. Przy doborze uwzględniono również cel i przedmiot badań, jak również jednostkę czasu [Grabiński i in. 1989]. Pomimo iż na stopień rozwoju społeczno-gospodarczego wpływają zjawiska o charakterze ilościowym i jakościowym, w badaniach skupiono się jedynie na zmiennych ilościowych, tak aby uzyskać możliwie wysoką obiektywność otrzymanych wyników i wniosków. Do analiz przyjęto zmienne o charakterze względnym (wskaźniki

struktury i natężenia) [Strahl 1998], pozyskane z Banku Danych Lokalnych Głównego Urzędu Statystycznego dla poziomu powiatów (NTS-4) w 2014 r. Do analiz, opierając się na literaturze przedmiotu, przyjęto zestaw 13 zmiennych, opisujących warunki społeczno-gospodarcze powiatów. Zmienne opisują warunki demograficzne, ekonomiczne oraz gospodarcze (tab. 1).

Tabela 1. Przyjęty do analizy wyjściowy zestaw zmiennych
Table 1. The starting data set accepted for the analysis

Nr	Nazwa zmiennej – Name of variable
x_1	Przyrost naturalny na 1000 mieszkańców – Population growth per 1000 inhabitants
x_2	Gęstość zaludnienia, os. · km ⁻² – Population density [person · km ⁻²
x_3	Odsetek ludności w wieku produkcyjnym, % – Percentage of the working-age population, %
x_4	Bezrobocie rejestrowane, % – Registered unemployment, %
x_5	Procentowy udział osób pracujących w sektorze usług w ogólnej liczbie ludności, % Percent share of people working in the services sector in total number of population, %
x_6	Zatrudnienie przeciętne [osób] – Average employment [persons]
x_7	Liczba podmiotów gospodarki narodowej na 10 tys. ludności Number of entities of national economy per 10 thousand of population
x_8	Udział nowo-zarejestrowanych podmiotów sektora kreatywnego w liczbie nowo zarejestrowanych podmiotów ogółem Share of newly-registered entities of creative sector in the number of newly-registered entities altogether
x_9	Udział nowo zarejestrowanych podmiotów sektora przetwórstwa rolno-spożywczego w ogólnej liczbie nowo zarejestrowanych podmiotów ogółem Share of newly-registered entities of agricultural and food processing sector in total number of newly-registered entities
x_{10}	Liczba ludności korzystającej z instalacji wodociągowej, osób Number of people using water supply system, persons
x_{11}	Liczba ludności korzystającej z instalacji kanalizacyjnej, osób Number of people using waste water system, persons
x_{12}	Udział obszarów prawnie chronionych według BDOT10k, % Share of legally protected areas according to BDOT10k, %
x_{13}	Ilość odpadów zebranych z gospodarstw domowego na 1 mieszkańca, kg Amount of wastes collected from households per 1 inhabitant, kg

W przeprowadzonych analizach podstawę klasyfikacji i wyodrębniania skupień obiektów w n -elementowym zbiorze obiektów w m -wymiarowej przestrzeni cech stanowiła macierz odległości D [$n \times n$] o d_{ij} elementach d_{ij} , które odpowiadały odległościom pomiędzy parami obiektów przyporządkowanych do i -tej kolumny oraz j -tego wiersza w unormowanej macierzy danych X [$n \times m$] [Murtagh i Contreras 2011]. Przyjęta definicja odległości czyli tzw. metryki mówi, że obiekt A będzie leżał bliżej obiektu B niż C (będzie bardziej podobny do obiektu B niż do obiektu C), jeśli odległość między

objektami A i B będzie mniejsza od odległości między obiektami A i C [Kolenda 2006]. Spośród metryk wybrano opartą na macyrcie Minkowskiego odległość euklidesową wykonując obliczenia według formuły (1):

$$d_{ik} = \sqrt[p]{\sum_{j=1}^m |x_{ij}' - x_{kj}'|^p} \quad (1)$$

gdzie:

$d_{i,k}$ – odległość metryczna pomiędzy i -tym a k -tym objektem,
 x_{ij}' , x_{kj}' – wartości w i -tym oraz k -tym wierszu unormowanej macierzy danych X'
 (tzn. wartości przyporządkowane i -temu oraz k -temu obiektowi),

$p = 2$.

Zastosowana do obliczeń metryka umożliwiła określenie homogeniczności oraz heterogeniczności skupień, stanowiła więc podstawę klasyfikacji powiatów woj. małopolskiego w trójwymiarowej przestrzeni zmiennych diagnostycznych opisujących zjawisko rozwoju społeczno-gospodarczego.

Do analiz porównawczych wybrano spośród metod hierarchicznych: metodę Czekanowskiego, aglomeracyjną kombinatoryczną metodę Warda oraz metodę taksonomii przestrzennej.

Metoda Czekanowskiego jest uznawana za pierwszą metodę taksonomii numerycznej. Jest często stosowana w szerokim spektrum badań. Polega na skonstruowaniu diagramu na podstawie macierzy odległości. W metodzie Czekanowskiego – odległość miejska $d(x,y)$ jest wartością bezwzględną różnicy i -tych wartości dla dwóch badanych obiektów np. x i y , jak w formule (2):

$$d(x, y) = \sum_{i=0}^n |x_i - y_i| \quad (2)$$

Pierwotnie diagram ten miał charakter asymetryczny, jednak obecnie stosuje się wersję przedziałową, czyli symetryczną, w której najmniejsze wartości (najbardziej podobne objekty) znajdują się blisko przekątnej [Panek 2009]. Obiektom należącym do określonych klas przyporządkowuje się symbole tworzące diagram. Diagram Czekanowskiego wymaga uporządkowania, co umożliwiają techniki komputerowe i zastosowanie odpowiednich algorytmów [Jaskulski i Sołtysiak 2004].

W metodzie Warda (aglomeracyjnej kombinatorycznej) każdy obiekt stanowi początkowo odrębną grupę, następnie zmniejsza się sukcesywnie liczbę grup przez ich łączenie w grupy wyższego rzędu. Proces łączenia kończy się z chwilą uzyskania jednej grupy składającej się ze zbioru wszystkich obiektów. Ogólna formuła metod aglomeracyjnych jest opisywana za pomocą schematu, nazywanego centralną procedurą aglomeracyjną [Lance i Williams 1967] (3):

$$d_{ir} = \alpha_p \cdot d_{ip} + \alpha_q \cdot d_{iq} + \beta \cdot d_{pq} + \gamma \cdot |d_{ip} - d_{iq}| \quad (3)$$

gdzie:

d_{ir} , d_{ip} , d_{iq} , d_{pq} – odległości między poszczególnymi grupami,
 α_p , α_q , β , γ – parametry charakterystyczne dla metod aglomeracyjnych.

Tabela 2. Macierz odległości euklidesowych obliczona dla przyjętego zestawu 13 zmiennych wyjściowych dla przeprowadzonych klasyfikacji
 Table 2. The Euclidean distance matrix calculated for the accepted set of 13 starting variables for performed classifications

Powiaty district	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	0,000	1,910	5,010	1,842	3,853	1,295	9,999	1,717	1,607	1,301	2,179	3,812	3,981	2,861	1,294	1,985	3,425	1,301	3,608	0,861	2,624	1,749
2	1,910	0,000	3,502	2,029	2,422	1,396	8,852	2,286	2,838	3,076	1,112	2,422	2,752	1,425	0,939	1,439	2,052	0,839	1,859	1,236	1,329	1,828
3	5,010	3,502	0,000	4,706	1,374	4,334	10,271	4,639	5,923	6,303	3,539	1,592	2,166	2,272	4,113	3,675	1,868	3,738	1,915	4,390	2,512	3,919
4	1,842	2,029	4,706	0,000	3,348	0,751	9,862	3,318	1,406	2,682	1,284	4,025	4,495	3,057	1,177	1,046	2,849	1,731	3,093	1,254	2,255	2,989
5	3,853	2,422	1,374	3,348	0,000	3,032	10,027	3,856	4,605	5,151	2,227	1,550	2,310	1,508	2,869	2,325	0,499	2,584	0,993	3,164	1,245	3,112
6	1,295	1,396	4,334	0,751	3,032	0,000	9,681	2,607	1,603	2,382	1,012	3,457	3,857	2,458	0,480	0,820	2,548	1,000	2,744	0,540	1,823	2,265
7	9,999	8,852	10,271	9,862	10,027	9,681	0,000	10,184	10,166	10,248	9,293	9,807	9,840	9,496	9,349	9,858	9,939	9,625	9,053	9,608	9,748	10,161
8	1,717	2,286	4,639	3,318	3,856	2,607	10,184	0,000	3,320	2,427	3,064	3,150	2,997	2,502	2,361	2,965	3,593	1,845	3,711	2,103	2,878	0,757
9	1,607	2,838	5,923	1,406	4,605	1,603	10,166	3,320	0,000	1,664	2,526	5,006	5,328	4,005	1,935	2,298	4,113	2,408	4,325	1,699	3,413	3,301
10	1,301	3,076	6,303	2,682	5,151	2,382	10,248	2,427	1,664	0,000	3,339	5,048	5,134	4,115	2,451	3,161	4,717	2,586	4,859	2,052	3,916	2,784
11	2,179	1,112	3,539	1,284	2,227	1,012	9,293	3,064	2,526	3,339	0,000	2,888	3,418	1,951	0,919	0,588	1,747	1,225	1,824	1,323	1,158	2,542
12	3,812	2,422	1,592	4,025	1,550	3,457	9,807	3,150	5,006	5,048	2,888	0,000	0,762	1,021	3,140	3,028	1,750	2,621	1,687	3,337	1,840	2,484
13	3,981	2,752	2,166	4,495	2,310	3,857	9,840	2,997	5,328	5,134	3,418	0,762	0,000	1,468	3,503	3,551	2,483	2,924	2,366	3,639	2,442	2,443
14	2,861	1,425	2,272	3,057	1,508	2,458	9,496	2,502	4,005	4,115	1,951	1,021	1,468	0,000	2,123	2,108	1,388	1,626	1,339	2,330	1,043	1,810
15	1,294	0,939	4,113	1,177	2,869	0,480	9,349	2,361	1,935	2,451	0,919	3,140	3,503	2,123	0,000	0,955	2,409	0,694	2,481	0,441	1,641	2,004

16	1,985	1,439	3,675	1,046	2,325	0,820	9,858	2,965	2,298	3,161	0,588	3,028	3,551	2,108	0,955	0,000	1,828	1,169	2,164	1,208	1,222	2,437
17	3,425	2,052	1,868	2,849	0,499	2,548	9,939	3,593	4,113	4,717	1,747	1,750	2,483	1,388	2,409	1,828	0,000	2,174	0,927	2,711	0,813	2,860
18	1,301	0,839	3,738	1,731	2,584	1,000	9,625	1,845	2,408	2,586	1,225	2,621	2,924	1,626	0,694	1,169	2,174	0,000	2,324	0,725	1,361	1,362
19	3,608	1,859	1,915	3,093	0,993	2,744	9,053	3,711	4,325	4,859	1,824	1,687	2,366	1,339	2,481	2,164	0,927	2,324	0,000	2,851	1,191	3,038
20	0,861	1,236	4,390	1,254	3,164	0,540	9,608	2,103	1,699	2,052	1,323	3,337	3,639	2,330	0,441	1,208	2,711	0,725	2,851	0,000	1,920	1,838
21	2,624	1,329	2,512	2,255	1,245	1,823	9,748	2,878	3,413	3,916	1,158	1,840	2,442	1,043	1,641	1,222	0,813	1,361	1,191	1,920	0,000	2,177
22	1,749	1,828	3,919	2,989	3,112	2,265	10,161	0,757	3,301	2,784	2,542	2,484	2,443	1,810	2,004	2,437	2,860	1,362	3,038	1,838	2,177	0,000

Cyframi oznaczono powiaty – Districts were labeled with numbers: 1 – bocheński, 2 – krakowski, 3 – miechowski, 4 – myślenicki, 5 – proszowicki, 6 – wielicki, 7 – miasto Kraków, 8 – gorlicki, 9 – limanowski, 10 – nowosądecki, 11 – miasto Nowy Sącz, 12 – chrzanowski, 13 – olkusi, 14 – oświęcimski, 15 – wadowicki, 16 – brzeski, 17 – dąbrowski, 18 – tarnowski, 19 – miasto Tarnów, 20 – nowotarski, 21 – suski, 22 – tatrzański

Metoda ta jest jedną z najczęściej stosowanych. Odległość pomiędzy elementami (grupami) jest definiowana jako różnica sum kwadratów odchyłeń odległości poszczególnych elementów od środków ciężkości skupienia, do którego należą elementy [Grabinski 1992].

Metoda taksonomii przestrzennej jest jedną z metod obszarowych. Metody obszarowe stanowią najbardziej rozbudowaną grupę metod taksonomicznych. Ogólna zasada klasyfikacji polega na podziale wielowymiarowej przestrzeni, w której znajdują się grupowane obiekty, na podprzestrzenie za pomocą hiperkul lub hiperkostek. Promienie kul lub krawędzie kostek wyznacza się ogólnie, najczęściej na podstawie wzoru [Panek 2004] (4):

$$d_0 = \bar{d} + c \cdot s_d, \quad (4)$$

gdzie:

- \bar{d} – średnia arytmetyczna,
- s_d – odchylenie standardowe liczone jako minimalna odległość poszczególnych obiektów do obiektów pozostałych,
- c – liczba rzeczywista, wybrana dowolnie, poszukuje się jej w sposób empiryczny; może przyjmować wartości 2, -1, -1/2, 0, 1/2, 1 itp.

Metoda taksonomii przestrzennej polega na łączeniu hiperkul konstruowanych dla każdego punktu.

Oceny możliwości wykorzystania wyżej wymienionych metod do badania zjawisk społeczno-gospodarczych dokonano w oparciu o mierniki oceny klasyfikacji, które służą do oceny skonstruowanej klasyfikacji. Spośród mierników oceny klasyfikacji wybrano mierniki indywidualne, mierniki homogeniczności skupień oraz heterogeniczności, a także mierniki poprawności skupień [Kolenda 2006].

Mierniki indywidualne – oznaczone I1, I2, ..., I6 – są wyznaczane tylko dla jednej grupy obiektów i określają jej homogeniczność lub heterogeniczność. Mierniki I1, I2, I3, I4, I6 określają homogeniczność (zwartość) grupy. Im mniejsza wartość miernika, tym obiekty w grupie leżą bliżej siebie, co oznacza jednocześnie, że są bardziej do siebie podobne. Miernik I5 określa heterogeniczność (odrębność) grupy, według założenia, że im większa wartość miernika, tym grupa jest bardziej oddalona od pozostałych grup.

Mierniki homogeniczności skupień są oznaczane jako M1, M2, ..., M18, natomiast mierniki heterogeniczności skupień jako M19, M20, ..., M30. Dodatkowo wyróżnia się mierniki poprawności skupień – M31, M32, ..., M45, a także mierniki pozostałe – D, Dm, H, h, SC, SK, V (wariancja wewnątrzgrupowa). Średnica taksonomiczna D, obliczana dla całego zbioru obiektów, oznacza maksymalną odległość w macierzy odległości, na podstawie której badane jest podobieństwo obiektów. Miernik Dm to średnia wartość odległości w macierzy odległości. Mierniki typu h (H) – określające heterogeniczność grupy według kryteriów:

- 1) h** – bardzo dobra heterogeniczność grupy,
- 2) h* – dobra heterogeniczność grupy,
- 3) h – słaba lecz dopuszczalna heterogeniczność grupy,
- 4) „-” – brak heterogeniczności grupy.

Dla każdego obiektu wyznaczana jest pomocnicza wartość „ s_i ” według formuły (5):

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (5)$$

gdzie:

a_i – średnia odległość obiektu i względem wszystkich obiektów w swojej grupie,
 b_i – średnia odległość obiektu i względem wszystkich obiektów w grupie sąsiedniej, w której znajduje się obiekt będący najbliższym sąsiadem dla obiektu i spośród obiektów spoza swojej grupy.

$s_i = 0$ dla grupy zawierającej jeden punkt. SK oznacza średnią wartość s_i dla wszystkich punktów należących do grupy. Natomiast miernik SC to maksymalna wartość SK [Kolenda 2006]. Miernik SC – *silhouette coefficient* według Rousseeuwa [1987] oznacza przy określonych wartościach podział na grupy:

- 0,71 – 1,00 – bardzo dobry,
- 0,51 – 0,70 – dobry,
- 0,26 – 0,50 – słaby, wraz z zaleceniem zastosowania innej metody klasyfikacji,
- 0,25-0,00 – nieodpowiedni, a wyodrębnione grupy nie tworzą żadnych grup obiektów podobnych.

WYNIKI BADAŃ I POCZYNIONE OBSERWACJE

Przestrzenne porównanie wyników analiz (tab. 3) wskazuje na różne wyniki klasyfikacji, w zależności od przyjętych metod klasyfikacyjnych. Dla każdej z metod sprawdzono wyniki delimitacji przy podziale na 5 grup obiektów. Porównywalne wyniki klasyfikacji otrzymano dla wydzielonych skrajnych grup, czyli obiektów charakteryzujących się z jednej strony najlepszymi, a z drugiej najsłabszymi własnościami. Najwyżej w każdej z przeprowadzonych klasyfikacji oceniono miasta na prawach powiatu: Kraków, Tarnów oraz Nowy Sącz.

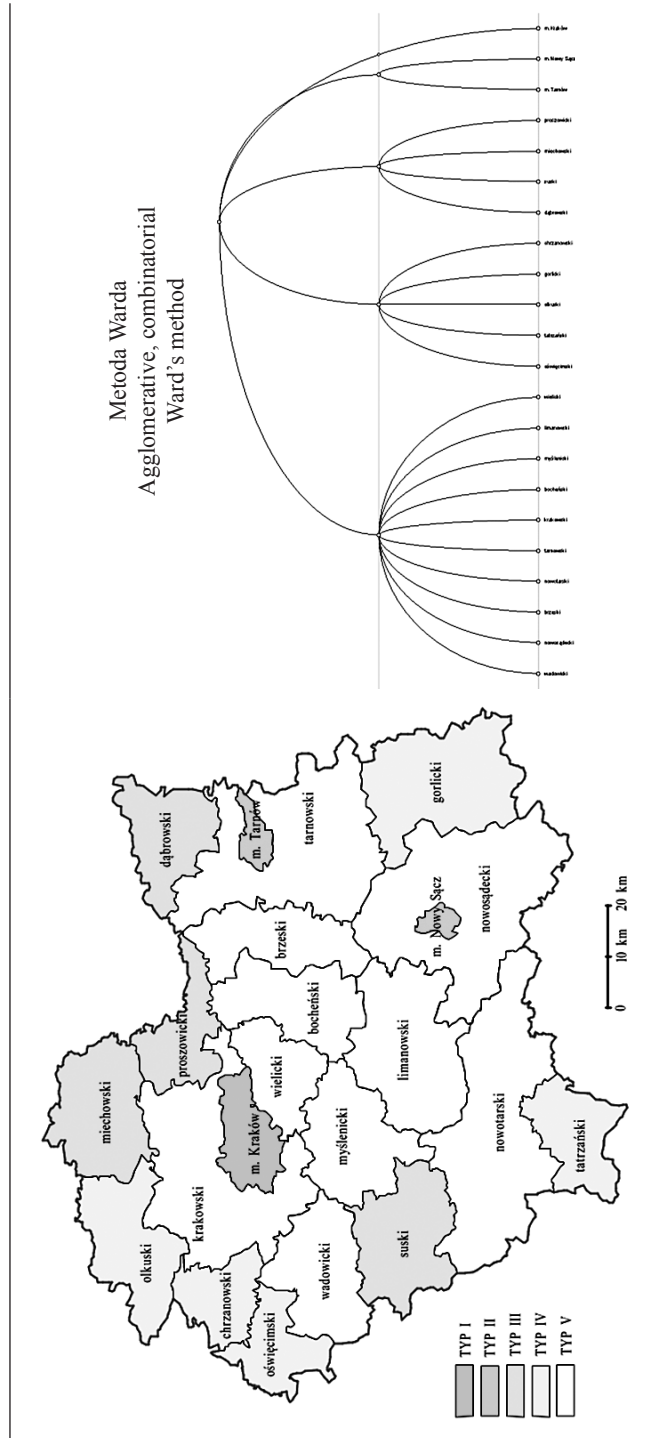
Zastosowane metody taksonomiczne doprowadziły do wydzielenia grup obiektów różniących się pomiędzy sobą liczebnością. Największe dysproporcje w liczebności grup daje metoda taksonomii przestrzennej, najmniejsze metoda Warda. Metoda Czekanowskiego doprowadziła do wydzielenia dwóch grup obiektów o liczebności 9 i 8 obiektów oraz dwóch grup po 2 obiekty, pozostawiając miasto Kraków w jednoobiektowej grupie (tab. 3, tab. 5).

Porównanie mierników oceny klasyfikacji (tab. 4) wykazało najlepsze dopasowanie przeprowadzonej klasyfikacji w przypadku grupowych mierników homogeniczności dla aglomeracyjnej metody Warda. Można zatem potwierdzić wysoką ocenę tej grupy metod aglomeracyjnych kombinatorycznych przytaczaną przez Grabińskiego [1989].

Najniższe wartości mierników homogeniczności (M2, M4, M14) odnotowano w przypadku metody Warda, co świadczy o największej wartości wydzielonych w procesie klasyfikacji grup. Drugie miejsce w rankingu pod względem homogeniczności zajmuje metoda Czekanowskiego. Mierniki heterogeniczności wykazują najlepszy wynik dla metody taksonomii przestrzennej. Wartości mierników poprawności skupień wskazują zarówno na metodę Warda jak i na taksonomię przestrzenną. Wartości miernika SC –

Tabela 3. Porównanie wyników klasyfikacji metodą Warda, Czekanowskiego oraz taksonomii przestrzennej. Przestrzenne rozmieszczenie wyników klasyfikacji zestawiono z wizualizacją w postaci grafów

Table 3. Comparison of classification results of Ward's and Czekanowski's methods as well as spatial taxonomy. Spatial location of classification results was compared with visualization in the form of graphs



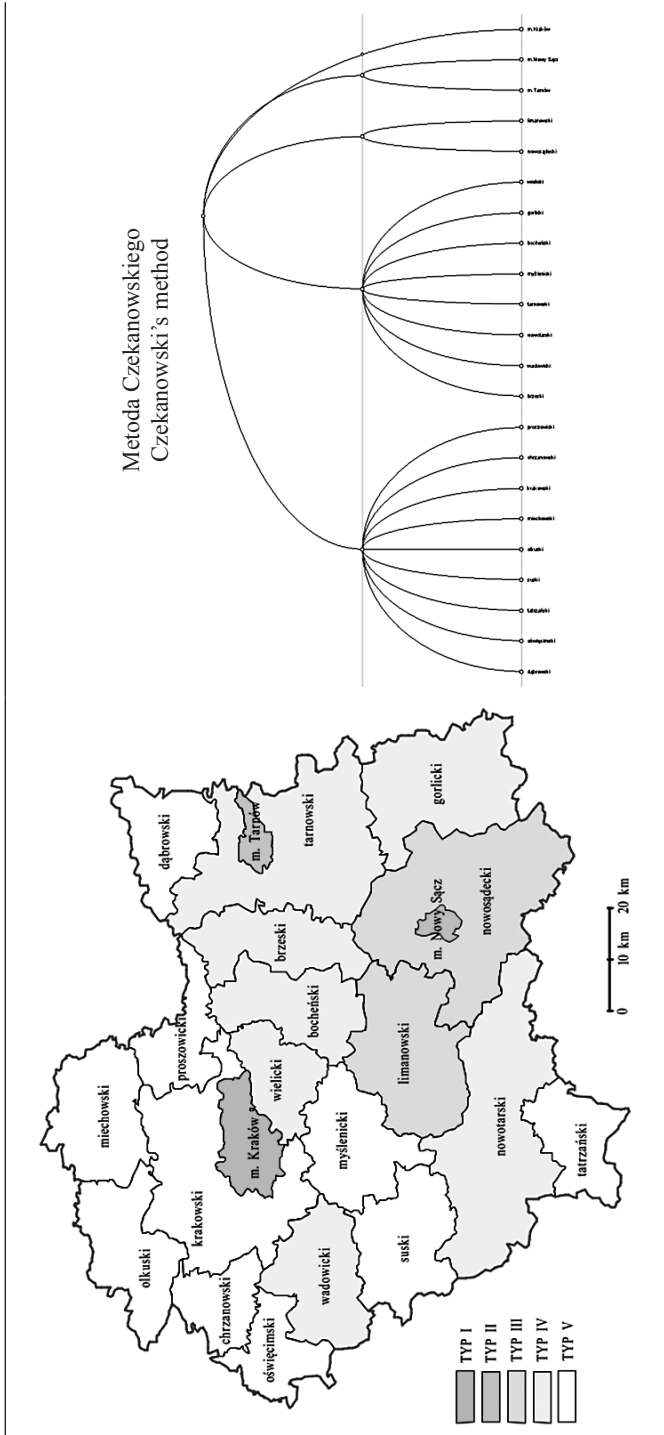
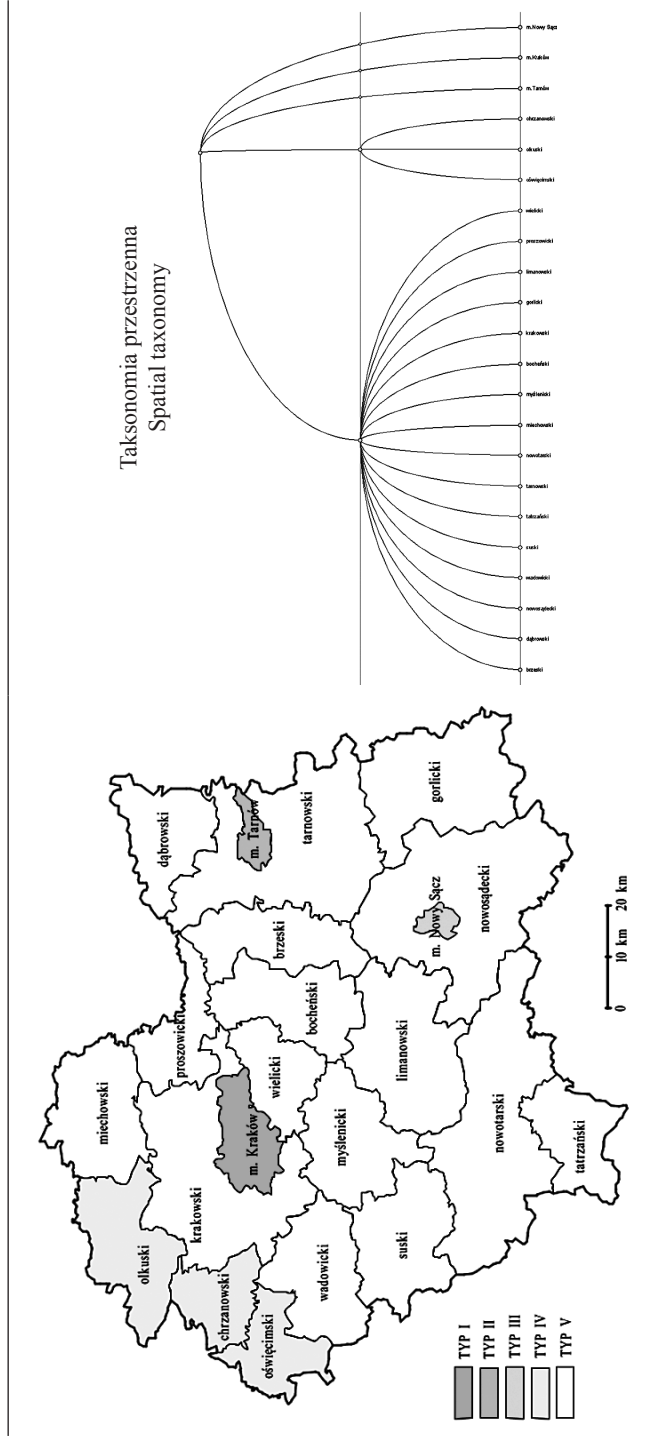


Tabela 3. cd.
Table 3. cont.



Źródło: opracowanie własne.
Source: own study.

silhouette coefficient różnią się nieznacznie pomiędzy dwoma pozostałymi metodami, jednak dla metody taksonomii przestrzennej oznaczają dopasowanie na poziomie dobrym, dla metody Warda tylko słabym (0,499), chociaż można uznać ten wskaźnik za wartość na granicy błędu obliczeń [Rousseeuw 1987]. Najmniejszą wariację wewnątrzgrupową uzyskano stosując aglomeracyjną metodę Warda. Dla każdej z metod maksymalna wartość odległości w macierzy euklidesowej wyniosła 15,358 jednostek, przy średniej wartości odległości równej 4,180 jednostek. Należy jednak pamiętać, iż mierniki oceny klasyfikacji każdorazowo zależą od liczby wydzielanych grup, liczby przyjętych do analiz zmiennych oraz od konfiguracji wartości tych zmiennych.

Tabela 4. Porównanie wybranych mierników oceny klasyfikacji uzyskanych dla metody taksonomii przestrzennej, Warda oraz Czekanowskiego

Tabela 4. Comparison of selected standards of measurement to evaluate classifications obtained for spatial taxonomy and also Ward's and Czekanowski's methods

Mierniki oceny klasyfikacji (dla rzędów) Classification standards of measurement	Taksonomia przestrzenna Spatial taxonomy	Aglomeracyjna, kombinatoryczna metoda Warda Agglomerative, combinatorial Ward's method	Metoda Czekanowskiego Czekanowski's method
D		15,358	
Dm		4,180	
H	–	–	–
K (liczebność) K (number)		5	
M2	2,713	2,386	2,563
M4	6,412	3,628	4,248
M14	4,232	2,830	3,018
M20	12,573	12,573	12,573
M21	2,045	1,315	1,315
M32	1,326	1,814	1,949
M34	3,136	2,759	3,230
SC	0,520	0,499	0,488
V	0,235	0,137	0,155

Porównanie indywidualnych mierników oceny klasyfikacji (tab. 5) wykazuje nieznaczne różnice pomiędzy poszczególnymi metodami klasyfikacji. Najlepsze dopasowanie miernika indywidualnego dla grup II występuje w przypadku metody Warda, według miernika I2 najlepiej sklasyfikowano obiekty metodą Czekanowskiego, natomiast miernik I5 wskazuje na metodę taksonomii przestrzennej. Miernik SK najwyższą wartość przyjmuje również dla taksonomii przestrzennej, przy nieznacznie niższych wartościach dla pozostałych dwóch metod.

Tabela 5. Indywidualne mierniki oceny klasyfikacji obliczone dla wydzielonych grup obiektów
 Table 5. Individual standards of measurement for evaluating classifications calculated for selected groups of objects

Mierniki dla grup	Taksonomia przestrzenna Spatial taxonomy					Aglomeracyjna, kombinatoryczna met. Warda Agglomerative, combinatorial Ward's method					Mierniki dla grup					Metoda Czekanowskiego Czekanowski's method				
	grupa 1	grupa 2	grupa 3	grupa 4	grupa 5	grupa 1	grupa 2	grupa 3	grupa 4	grupa 5	grupa 1	grupa 2	grupa 3	grupa 4	grupa 5	grupa 1	grupa 2	grupa 3	grupa 4	grupa 5
II	2,7	1,6	j.o.	j.o.	j.o.	II	2,0	2,3	1,8	2,1	j.o.	II	2,7	1,7	2,0	2,1	j.o.			
I2	6,4	2,2	j.o.	j.o.	j.o.	I2	3,6	3,5	2,7	2,1	j.o.	I2	4,2	3,5	2,0	2,1	j.o.			
I5	2,0	2,0	2,1	12,6	2,1	I5	1,3	2,0	1,3	3,4	12,6	I5	1,4	1,3	1,6	3,4	12,6			
K	16	3	1	1	1	K	10	5	4	2	1	K	9	8	2	2	1			
SK	0,2	0,52	0	0	0	SK	0,4	0,1	0,4	0,49	0	SK	0,1	0,4	0,2	0,49	0			
h	-	h	-	h**	-	h	-	-	-	h*	h**	h	-	-	-	h*/h'	h**			

j.o. – jedna obserwacja (jeden obiekt) – one observation (one object), K – liczebność grup – numerical amount of groups

PODSUMOWANIE I WNIOSKI

Województwo małopolskie charakteryzuje się dużymi dysproporcjami rozwoju społeczno-gospodarczego. Analiza przeprowadzona wśród małopolskich powiatów potwierdziła badania autorów i znany osąd o słabiej rozwiniętej części południowo-wschodniej oraz wyróżniającej się dodatnio części północnej oraz zachodniej [Prus i Szylar 2015; Moskal 1997, Kotala 1998]. Wśród powiatów woj. małopolskiego wyróżniają się w szczególności powiaty miejskie: krakowski, tarnowski oraz nowosądecki, charakteryzujące się wysokim poziomem rozwoju na tle powiatów miejsko-wiejskich. Zauważalny jest także proces polaryzacji przestrzennej tzn. powiększają się różnice rozwojowe pomiędzy obszarami lepiej rozwiniętymi i obszarami biednymi (problemowymi) [Bański 2009, Bański i Czapiewski 2008, Rosner 2010]. Regionalne zróżnicowanie rozwoju według Myrdala [1957] wynika z uwarunkowań historycznych, w tym ekonomicznych, społecznych i kulturowych, oraz z czasem ulega pogłębieniu tj. szybszego rozwoju obszarów zamożniejszych przy zacofaniu regionów słabszych. Można zaobserwować również, że tempo rozwoju społeczno-gospodarczego powiatów woj. małopolskiego jest nierównomierne i skupia się w centrach wzrostu, co uzasadnia m.in. teorię Hirschmana [1958].

Rozwój społeczno-gospodarczy jest zjawiskiem definiowanym w literaturze przedmiotu w wielowymiarowej przestrzeni [Klóska 2012, Ziemiańczyk 2010, Gawroński i in. 2014], co ze statystycznego punktu widzenia utrudnia jego jednoznaczny sposób pomiaru. Wyniki badań zależą od listy wybranych do analizy zmiennych diagnostycznych, od wyboru metody normalizacji zmiennych, doboru współczynników wagowych, doboru miary podobieństwa pomiędzy obiektami (odległości taksonomicznej), jak również metody grupowania. Jak pokazały badania, zastosowanie trzech różnych metod klasyfikacji obiektów tj. metody taksonomii przestrzennej, metody Warda oraz Czekanowskiego do badania zjawiska rozwoju społeczno-gospodarczego, przy uwzględnieniu jednolitego zestawu zmiennych diagnostycznych, jednakowej metodzie normalizacji oraz w oparciu o jednakową metrykę (odległość taksonomiczną) doprowadziło również do otrzymania różniących się od siebie wyników. Potwierdza to pogląd, iż różne formuły agregacji dają różne końcowe wyniki, nawet przy zastosowaniu jednakowej listy zmiennych diagnostycznych [Klóska 2012, Kolenda 2006]. Niemniej jednak można odnaleźć pewne cechy wspólne zastosowanych metod klasyfikacji oraz podobieństwo w zakresie wydzielonych grup obiektów homogenicznych. Można także zaobserwować pewne prawidłowości. Metoda taksonomii przestrzennej prowadzi do wydzielenia kilku grup złożonych z pojedynczych obiektów oraz jednej grupy skupiającej ponad 70% jednostek poddanych analizie. W aglomeracyjnej kombinatorycznej metodzie Warda oraz w metodzie Czekanowskiego rozkład jednostek w wydzielonych grupach jednorodnych jest bardziej równomierny niż w przypadku taksonomii przestrzennej. Ocena i wybór najlepszej z metod klasyfikacji obiektów jest relatywnie trudna, ponieważ jak wykazują badania, poprawność wyników grupowania zależy m.in. od struktury punktów w przestrzeni wielowymiarowej i od rozkładu cech, a także od wyboru kryterium oceny klasyfikacji [Kolenda 2006]. W oparciu o mierniki oceny klasyfikacji można dokonać wyboru najlepszej metody klasyfikacji w przypadku wielowymiarowych zjawisk społeczno-gospodarczych. Klasyfikacja pod kątem statystycznym wykazuje najlepsze wskaźniki świadczące

o homogeniczności w grupach w przypadku metody Warda. Warunek heterogeniczności jest najlepiej spełniony w przypadku metody taksonomii przestrzennej. Najslabsze mierniki klasyfikacji, zarówno te świadczące o homogeniczności, jak i heterogeniczności, zaobserwowano po zastosowaniu metody Czekanowskiego. Przeprowadzone badania potwierdzają najwyższą skuteczność w przypadku metody aglomeracyjnej kombinatorycznej Warda. Pozwala to na stwierdzenie, iż wśród metod taksonomicznych lepsze wyniki dają metody aglomeracyjne, nie zaś podziałowe. Wyróżniają się również zalety metod wzorcowych i hierarchicznych – ukazujących w przejrzysty sposób strukturę podziału badanego zjawiska. Grabiński i in. [1989] ustalili m.in., że nie ma zastrzeżeń co do wyników grupowania metodami należącymi do grupy kombinatorycznych aglomeracyjnych oraz do metod optymalizacyjnych, pod warunkiem, że wstępny podział ustalany był za pomocą aglomeracyjnych metod kombinatorycznych. M. Kolenda [2006] stwierdziła natomiast, iż dobre wyniki są możliwe do osiągnięcia przy zastosowaniu metody taksonomii przestrzennej, jednakże pod warunkiem standardowo ustalanych parametrów (niekiedy lepsze wyniki od metody Warda), czego nie potwierdziły prowadzone w niniejszym artykule badania.

LITERATURA

- Ahavar, E., Daneshgar-Moghaddam, N., Ortiz, A. M., Crespi, N. (2016). On analyzing user location discovery methods in smart homes: A taxonomy and survey. *J. Network and Computer Applic.*, 76, 75–86.
- Assunção, R. M., Neves, M. C., Câmara, G., da Costa Freitas, C. (2006). Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *Internat. J. Geographical Information Sci.*, 20(7), 797–811.
- Bagdziński, S.L. (1994). *Lokalna polityka gospodarcza (w okresie transformacji systemowej)*. Wyd. Uniwersytetu Mikołaja Kopernika, Toruń.
- Bański, J. (red.) (2009). *Analiza zróżnicowania i perspektyw rozwoju obszarów wiejskich w Polsce do 2015 roku. Studia Obszarów Wiejskich, XVI*. Komisja Obszarów Wiejskich PTG, Warszawa.
- Bański, J., Czapiewski, K. (2008). *Identyfikacja i ocena czynników sukcesu społeczno-gospodarczego na obszarach wiejskich*. Instytut Geografii i Przestrzennego Zagospodarowania PAN, Warszawa.
- Bolis, I., Morioka, S., Szelwar, L. (2017). Are we making decisions in a sustainable way? A comprehensive literature review about rationalities for sustainable development. *J. Cleaner Production*, 14, 310–322.
- Cieszewska, A. (2008). Zachowanie terenów cennych przyrodniczo w kształtowaniu struktury krajobrazu na poziomie miejscowego planu zagospodarowania przestrzennego. *Struktura i funkcjonowanie systemów krajobrazowych: meta-analzy, modele, teorie i ich zastosowania. Problemy ekologii krajobrazu*, 21, 239–250.
- Clifford, H.T., Stephenson, W. (1975). *An introduction to numerical classification*. New York Academic Press INC.
- Chojnicki, Z., Czyż, T. (1973). *Metody taksonomii numerycznej w regionalizacji geograficznej*. Państwowe Wydawnictwo Naukowe, Warszawa.
- Czekanowski, J. (1913). *Zarys metod statystycznych w zastosowaniu do antropologii*. Prace Towarzystwa Naukowego Warszawskiego, 5.
- Dunn, G., Everitt, B.S. (2004). *An introduction to Mathematical Taxonomy*. Dover Publications, Mineola, NY.

- Gawroński, K., Prus, B., Sołtysik, S. (2014). Analiza i ocena warunków rozwoju społeczno-gospodarczego województwa podkarpackiego. *Infrastr. Ekol. Ter. Wiej.*, 4(2), 1241–1254.
- Fierich, J. (1957). Próba zastosowania metod taksonomicznych do rejonizacji systemów rolniczych w województwie krakowskim. *Myśl Gospodarcza*, 3.
- Frankowski, Z. (1991). Zastosowanie metod taksonomicznych w badaniach przestrzennych. Agencja Wydawnicza Instytutu Gospodarki Przestrzennej i Komunalnej, Warszawa.
- Gorzelać, G. (1992). Planowanie rozwoju lokalnego (omówienie książki E. Blakley-a: *Planning Local Economic Development*). *Studia Regionalne i Lokalne*, 7(40), 18–58.
- Grabiński, T., Wydymus, S., Zeliaś, A. (1989). *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*. PWN, Warszawa.
- Grabiński, T. (1992). *Metody aksometrii. Prace Pomocy Naukowo-dydaktycznych Akademii Ekonomicznej w Krakowie*, Kraków.
- Gutiérrez, E.E., Pine R.H. (2017). Specimen collection crucial to taxonomy. *Science*, 355(6331), 1275.
- Gwiaździńska-Goraj, M., Goraj, S. (2013). The Contribution of the Natural Environment to Sustainable Development on the Example of Rural Areas in the Region of Warmia and Mazury. *J. Rural Developm.*, 6, 266–271.
- Hałasiewicz, A. (2010). Rozwój obszarów wiejskich w kontekście zróżnicowań przestrzennych w Polsce i budowania spójności terytorialnej kraju. Ekspertyza wykonana na zlecenie MRR, *Warszawa*.
- Hellwig, Z. (1981). *Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych*. Wyd. PWE, Warszawa.
- Hełdak, M., Raszka, B. (2013). Evaluation of the local spatial policy in Poland with regard to sustainable development. *Pol. J. Environ. Stud.*, 22(2), 395–402.
- Hirschman, A. (1958). *The strategy of economic development*, Yale University Press, New Heaven.
- Jaskulski, P., Sołtysiak, A. (2004). Diagram Czekanowskiego: pomysły, historia, zastosowania. [W:] *Klasyfikacja i analiza danych – teoria i zastosowania*. Red. K. Jajuga, M. Walesiak. *Prace Naukowe Akademii Ekonomicznej im. Oskara Langego*, 1022, *Taksonomia*, 11, 374–383.
- Kates, R.W., Parris, T.M., Leiserowitz, A.A. (2005). What is sustainable development? Goals, Indicators, Values and Practice. *Environment: Sci. and Policy of Sustainable Develop.*, 47(3), 8–23.
- Kauffman, C., Martin, P. (2017). Can Rights of Nature Make Development More Sustainable? Why Some Ecuadorian lawsuits Succeed and Others Fail. *World Develop.*, 92, 130–142.
- Kotala, A. (1998). *Problemy i perspektywy przebudowy wsi i rolnictwa w Małopolsce*. Wydawnictwo Akademii Rolniczej w Krakowie, Kraków.
- Klavans, R., Boyack, K.W. (2017). Which Type of Citation Analysis Generates the Most Accurate Taxonomy of Scientific and Technical Knowledge? *J. Assoc. Inform. Sci. and Technol.*, 68(4), 984–998.
- Kłóska, R. (2012). Statystyczna analiza poziomu rozwoju społeczno-gospodarczego w Polsce w ujęciu regionalnym. *Zeszyty Naukowe Wyższej Szkoły Bankowej w Poznaniu*, 42, 127–135.
- Kostrubiec, B. (1982). *Taksonomia numeryczna w badaniach geograficznych*. Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław.
- Król, K. (2015). Conception of a touristic map and nature protection forms created with use of open data sources and free software on a Grybów commune example. *Geomatics, Landmanagement and Landscape*, 4, 49–59.
- Kuciński, K. (red.) (2009). *Geografia ekonomiczna*. Oficyna Wolters Kluwer Polska, Warszawa.
- Lance, G.N., Williams, W.T. (1967). A General Theory of Classificatory Sorting Strategies I. Hierarchical Systems. *The Computer J.*, 9(4), 373–380.

- Moskal, S. (1997). Założenia Małopolskiego Programu Rozwoju Wsi i Rolnictwa. *Studia Małopolskie*, 1, 85–106.
- Murtagh, F., Contreras, P. (2011). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1), 86–97.
- Myrdal, G. (1957). *Economic theory and underdeveloped regions*, Duckworth, London.
- Our Common Future (1987). World Commission on Environment and Development (WCED). Oxford University Press, New York.
- Panek, T. (2009). *Statystyczne metody wielowymiarowej analizy porównawczej*. Szkoła Główna Handlowa, Warszawa.
- Prus, B., Szylar, M. (2015). Evaluation of possibility to use typological procedure and Wrocław taxonomy to analyse agriculture development conditions on the example of the former Kraków voivodship. *Geomatics Landmanagement and Landscape*, 4, 83–101.
- Redclift, M. (2005). Sustainable development (1987–2005): an oxymoron comes of age. *Sustainable Develop.*, 13, 212–227.
- Rosner, A. (2010). Zróżnicowanie przestrzenne rozwoju społeczno-gospodarczego polskiej wsi i tendencje jego zmian. Kancelaria Senatu. Biuro Analiz i Dokumentacji, <http://www.senat.gov.pl/k7/dok/opinia/2010/oe-148.pdf> (dostęp: 20.05.2016).
- Rousseuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Computational and Appl. Mathematics*, 20, 53–65. doi: [http://dx.doi.org/10.1016/0377-0427\(87\)90125-7](http://dx.doi.org/10.1016/0377-0427(87)90125-7).
- Semenova, N., Busalova, S., Eremina, O., Makeikina, S., Ivanova, I. (2016). Assessment of Sustainable Development of Rural Areas of Russia. *Indian J. Sci. and Technol.*, 9(14), 1–6.
- Siekierski, J. (2005). Przemiany strukturalne w rolnictwie polskim w latach 1990-2003. *Prace Komisji Nauk Rolniczych i Biologicznych*, 57, 655–663.
- Sokołowski, D. (1999). Zróżnicowanie zbioru małych miast i większych osiedli wiejskich w Polsce w ujęciu koncepcji kontinuum wiejsko-miejskiego. Wydawnictwo UMK, Toruń.
- Strahl, D. (red.) (1998). *Taksonomia struktur w badaniach regionalnych*. Wydawnictwo Akademii Ekonomicznej, Wrocław.
- Szylar, M., Prus, B. (2015). Analiza stanu zagospodarowania przestrzennego przy wykorzystaniu metody Czekanowskiego. *Episteme*, 26(II), 141–149.
- Szymła, Z. (2000). *Determinanty rozwoju regionalnego*. Ossolineum, Wrocław – Warszawa – Kraków.
- Tavana, M., Liu, W., Elmore, P., Petry F.E., Bourgeois B.S. (2016). A practical taxonomy of methods and literature for managing uncertain spatial data in geographic information systems. *Measurement*, 81, 123–162.
- Ustawa z dnia 27 marca 2003 r. o planowaniu i zagospodarowaniu przestrzennym. Dz. U. z 2003 r. Nr 80, poz. 717 z późn. zmianami.
- Ustawa z dnia 27 kwietnia 2001 r. Prawo ochrony środowiska. Dz.U. z 2001 r. Nr 62, poz. 627 z późn. zmianami.
- Wilkin, J. (2005). *Lepszy świat – polska wieś za 25 lat. [w:] Polska wieś 2025*. Red. J. Wilkin. Fundusz Współpracy, Warszawa, 41–44.
- Ziemiańczyk, U. (2010). Ocena poziomu rozwoju społeczno-gospodarczego powiatów wiejskich i miejsko – wiejskich w województwie małopolskim. *Infrastr. Ekol. Ter. Wiej.*, 14, 31–40.
- Zimny, A. (2007). Społeczno-gospodarcze zróżnicowanie gmin regionu konińskiego. *Rocznik Koniński*, 16. Red. M. Szczepaniak. Wydawnictwo PWSZ, Konin 2007, 173–192.
- Zwick, H.S., Syed, S.A.S. (2017). The polarization impact of the crisis on the Eurozone labour markets: a hierarchical cluster analysis. *Appl. Economic Letters*, 24(7), 472–476.

EVALUATION OF USING SELECTED TAXONOMIC METHODS TO CLASSIFY SOCIO-ECONOMIC PHENOMENA

Abstract. Taxonomic methods are used to describe and classify complex socio-economic phenomena expressed by the numerous set of variables. These methods give different results despite the approved uniform set of initial research features. The aim of the paper was to present the results of classification of Małopolska province districts using three taxonomic methods and to evaluate the performed division. The set of variables defining conditions of socio-economic development was taken to the surveys. Three taxonomic methods – spatial taxonomy, Ward’s agglomerative method and Czekanowski’s method – were submitted for evaluation. Standards of measurement to evaluate units’ classification i.e. individual ones, of homogeneity and heterogeneity calculated for clusters separated as a result of the analysis and also standards of measurement of clusters’ correctness were used for evaluation. Calculations were performed in the programme *Numerical Taxonomy* (in Polish: *Taksonomia Numeryczna*). The descriptive-reference method as well as the logical analysis were used in the paper. The best indexes of homogeneity that evidence of correct classification’s performing were obtained in case of the Ward’s method. The condition of heterogeneity was best fulfilled in case of the method of spatial taxonomy. The weakest classification standards both telling of homogeneity and heterogeneity were obtained after using the Czekanowski’s method. Performed surveys confirmed the best efficiency in case of the Ward’s method.

Key words: spatial taxonomy, Ward’s method, Czekanowski’s method, socio-economic development, Euclidean distance

Zaakceptowano do druku – Accepted for print: 2.06.2017

Do cytowań – For citation: Prus, B., Król, K. (2017). Ocena zastosowania wybranych metod taksonomicznych do klasyfikacji zjawisk społeczno-gospodarczych. *Acta. Sci. Pol., Formatio Circumiectus*, 16(2), 179–197.