

**ZASTOSOWANIE METODY ANALIZY SKUPIEŃ
ORAZ WIELOWYMIAROWEJ ANALIZY
KORESPONDENCJI DO OCENY POZIOMU
INFRASTRUKTURY OBSZARÓW WIEJSKICH**

Irena Kropsz

Uniwersytet Przyrodniczy we Wrocławiu

Abstrakt. Wielofunkcyjny rozwój obszarów wiejskich wiąże się z pełnieniem przez nie większej liczby funkcji pozarolniczych. Poziom rozwoju infrastruktury odgrywa istotną rolę w procesie wielofunkcyjnego rozwoju obszarów wiejskich oraz w znacznym stopniu wpływa na poziom życia mieszkającej tam ludności. W opracowaniu dokonano analizy poziomu wyposażenia w podstawowe elementy infrastruktury technicznej, społecznej i ekonomicznej obszarów wiejskich na terenie wybranych gmin Dolnego Śląska. Następnie za pomocą metod statystycznych dokonano waloryzacji punktowej tych obszarów pod względem poziomu infrastruktury, wykazując zróżnicowanie pomiędzy gminami.

Słowa kluczowe: infrastruktura, obszar wiejski, zróżnicowanie, waloryzacja, Dolny Śląsk

WPROWADZENIE

Infrastruktura to czynnik aktywizujący postęp społeczno-gospodarczy będący jednocześnie jednym z ważnych wyznaczników warunków życia na wsi. O zasadności rozważań na temat infrastruktury przesądza także to, że jej poziom jest nadal niewystarczający i utrzymują się duże dysproporcje pomiędzy wyposażeniem w infrastrukturę wsi i miast [Pięcek 1997]. Lokowanie kapitału na obszarach wiejskich jest jednak uwarunkowane, jak wykazują doświadczenia państw Unii Europejskiej, dobrze rozwiniętą infrastrukturą [Wilkin 1996].

Dodatkową trudnością jest duże zróżnicowanie poziomu wyposażenia infrastrukturalnego. Obecne zróżnicowanie wyposażenia infrastrukturalnego wsi pogłębia również

dekapitalizacja obiektów infrastrukturalnych. Oznacza to niejednokrotnie konieczność poniesienia w przyszłości znacznych nakładów na samo odtworzenie substancji materialnej tych obiektów. Jest to wynik niezapewnienia w przeszłości nawet reprodukcji prostej, czego konsekwencją jest nie tylko obniżenie poziomu świadczonych usług infrastrukturalnych, ich gorsza jakość, lecz także istotne zmniejszenie się ich zakresu [Wilczyńska 1997].

CEL I ZAKRES BADAŃ

Doceniając ważną rolę infrastruktury w rozwoju wsi, podjęto badania dotyczące jej poziomu na obszarach wiejskich Dolnego Śląska. Celem badań było:

- wykazanie zróżnicowania obszarów wiejskich Dolnego Śląska pod względem wyposażenia infrastrukturalnego na podstawie waloryzacji punktowej,
- zastosowanie metody analizy skupień oraz wielowymiarowej analizy korespondencji do waloryzacji punktowej infrastruktury obszarów wiejskich Dolnego Śląska,
- określenie przydatności waloryzacji punktowej obszarów wiejskich pod względem ich wyposażenia w infrastrukturę w podejmowaniu decyzji.

W celu uzyskania materiałów źródłowych badaniem objęto wszystkie gminy wiejskie i obszary wiejskie gmin miejsko-wiejskich Dolnego Śląska. Zakres badań wybranych elementów infrastruktury obejmował analizę elementów infrastruktury technicznej, społecznej i ekonomicznej w 2004 roku¹. Uzyskane wyniki posłużyły do oceny stanu infrastruktury obszarów wiejskich oraz do waloryzacji obszarów wiejskich Dolnego Śląska z wykorzystaniem statystycznej metody analizy skupień i analizy korespondencji.

Dostępność poszczególnych elementów infrastruktury technicznej, takich jak: sieć wodociągowa, kanalizacyjna i gazowa określano wskaźnikiem liczby połączeń przypadających na 1000 mieszkańców oraz w kilometrach na 100 km², a długość dróg o nawierzchni twardej w kilometrach na 100 km² oraz jako procent dróg ogółem. Oczyszczalnie ścieków przeliczono na 100 km² oraz określono procent ścieków oczyszczanych. Elementy infrastruktury ekonomicznej (sklepy, targowiska stałe i sezonowe) przedstawiono, określając liczbę elementów infrastruktury przypadających na 1000 mieszkańców, a także na 100 km². Jednostki zarejestrowane w systemie REGON również przeliczono na 100 km² oraz na 10 tys. mieszkańców w wieku produkcyjnym. Wyposażenie w infrastrukturę społeczną przedstawiono, określając liczbę elementów infrastruktury przypadających na 1000 mieszkańców oraz uwzględniając przestrzenny charakter obszarów wiejskich, również na 100 km². Analizie poddano: przedszkola, szkoły podstawowe, gimnazja, apteki i punkty apteczne, przychodnie i ośrodki zdrowia oraz biblioteki.

¹ Badania były finansowane w ramach projektu badawczego KBN pt. „Procesy dostosowawcze obszarów wiejskich Dolnego Śląska do wymogów UE” nr 2 P06 R 0522, pod kierunkiem B. Kutkowskiej.

METODYKA BADAŃ

Do przeprowadzenia badań gminy zostały wybrane w sposób celowy, tak więc badania szczegółowe w 2004 roku dotyczyły 133 gmin. Do zbierania materiałów pierwotnych wykorzystano metodę statystyki powszechnej. Szczególnie przydatne okazały się informacje zawarte w ogólnodostępnej bazie danych GUS.gov.pl.

Do opracowania zebranych materiałów wykorzystano metodę opisową do interpretacji wyników [Kopeć 1983] i metodę porównawczą do porównania otrzymanych wyników w gminach [Kopeć 1983, Ryznar 1999, Stachak 2003, 2006]. Przy opracowaniu danych wykorzystano również dwie metody statystyczne:

1. Analizę skupień zastosowano do grupowania obiektów – gmin na podstawie wybranych elementów infrastruktury technicznej, społecznej i ekonomicznej.

Przed przystąpieniem do grupowania zmienne będące podstawą klasyfikacji były standaryzowane, co pozwoliło na klasyfikację obiektów pod względem zmiennych różnomianowych. Przez standaryzację wszystkie zmienne mają wariancję równą 1 i średnią arytmetyczną równą 0. Następnie, po utworzeniu szeregów statystycznych, wykonano analizę skupień. W opracowaniu zastosowano dwuetapowe podejście grupowania danych. Pierwszy krok polegał na wykorzystaniu metody aglomeracji w celu ustalenia liczby skupień na podstawie otrzymanego dendrogramu. W drugim kroku obiekty były grupowane metodą k -średnich, w której należy podać liczbę skupień, jaką chcemy uzyskać – w tym wypadku taką, którą uzyskano w wyniku aglomeracji.

W metodzie aglomeracji dla określenia odległości między skupieniami wykorzystano metodę „najbliższych sąsiadów” między skupieniami – tzw. metodę pojedynczego wiązania. Jako miarę odległości, które mogą występować pomiędzy poszczególnymi obiektami, w pracy zastosowano odległość euklidesową, czyli geometryczną w przestrzeni wielowymiarowej [Program Statistica 2000, Sagan 2001].

2. Wielowymiarową analizę korespondencji zastosowano do określenia waloryzacji infrastruktury obszarów wiejskich Dolnego Śląska. Waloryzacja polegała na nadaniu wartości punktowych poszczególnym rodzajom infrastruktury w zależności od poziomu ich wyposażenia na obszarach badanych gmin.

Wielowymiarowa analiza korespondencji to prosta analiza korespondencji prowadzona na macierzy kodów, gdzie poszczególne wiersze odpowiadają kolejnym obserwacjom, a kolumny – wariantom zmiennych. Dla każdej obserwacji (przypadku – gminy) zaznaczamy 1 w kategorii (wariancie cechy), do której dany przypadek należy, natomiast w pozostałych kategoriach stawiamy 0. Przy ocenie jakości rozwiązania o wybranej liczbie wymiarów pomocne są dodatkowe statystyki. Najbardziej istotnym elementem oceny otrzymanych wyników analizy korespondencji jest kwadrat cosinusa (jakość lub kwadrat korelacji z każdym wymiarem). Kolumna ta zawiera miary jakości dla każdego punktu w odniesieniu do każdego wymiaru. Suma wartości tych kolumn liczona po wszystkich wymiarach jest równa całkowitej jakości. Wartość ta może być również interpretowana jako korelacja danego punktu z odpowiednim wymiarem [Program Statistica 2000].

WYNIKI BADAŃ

W celu nadania punktacji określającej poziom wyposażenia w infrastrukturę na obszarach wiejskich w wybranych gminach położonych na Dolnym Śląsku wykorzystano analizę skupień, która pozwoliła na podział gmin na jednorodne skupienia o zbliżonym poziomie wyposażenia w dane elementy infrastruktury. W kolejnym kroku zastosowano wielowymiarową analizę korespondencji należącą do grupy wielowymiarowych metod skalowania. Podstawą do utworzenia macierzy kodów były gminy podzielone wcześniej za pomocą analizy skupień na jednorodne skupienia, charakteryzujące się podobnym wyposażeniem w wybrane elementy infrastruktury. Ze względu na poziom reprezentowany przez dane skupienie określano cztery kategorie dla cechy jakościowej poszczególnych rodzajów infrastruktury: najlepsze, lepsze, średnie i gorsze. Ostatecznie po weryfikacji uzyskanych wyników do dalszych analiz przyjęto trzy kategorie dla każdej analizowanej cechy jakościowej.

Następnie przeprowadzono wielowymiarową analizę korespondencji opartą na macierzy kodów. Macierz kodów odpowiadająca przykładowi Dolnego Śląska w opracowaniu zawierała 133 wiersze (przypadki – gminy obserwowane na obszarach wiejskich). Przypadki dla Dolnego Śląska rozpatrywano pod względem dziewięciu kategorii jakościowych, które określono wcześniej dla każdego rodzaju infrastruktury z wykorzystaniem analizy skupień. Dla każdej cechy jakościowej (dla infrastruktury technicznej, ekonomicznej oraz społecznej) wyodrębniono trzy kategorie wyłonione na podstawie trzech utworzonych skupień – jako infrastrukturę lepszą, średnią i gorszą. Przygotowane macierze kodowe poddano analizie w taki sposób, jakby to były zwykłe tablice dwudzielcze, a otrzymane wyniki dostarczyły współrzędnych kolumn, które pozwoliły na analizę relacji między poszczególnymi kategoriami, wynikającymi z długości pomiędzy poszczególnymi wierszami (przypadkami-gminami). Na zakwalifikowanie przypadków (gmin) do poszczególnych kategorii jakościowych miały wpływ częstości ich występowania. Rozkład profili częstości zaobserwowanych w gminach Dolnego Śląska dla wszystkich analizowanych kategorii jakościowych infrastruktury pokazano w tabeli 1. Dane przedstawiono jako udziały poszczególnych kategorii jakościowych jednej cechy – infrastruktury – w obrębie innej.

Z przedstawionych profili częstości wynika, że np. w tworzeniu grupy gmin o infrastrukturze technicznej w kategorii jakościowej lepszej, w porównaniu z infrastrukturą społeczną, gminy o wyposażeniu średnim stanowiły 43%, natomiast gminy gorzej wyposażone 22% populacji w tej grupie. Aż 62% wynosił udział gmin średnio wyposażonych w infrastrukturę ekonomiczną w obrębie kategorii gmin o infrastrukturze technicznej lepszej. W obrębie kategorii infrastruktury społecznej średniej gminy o wyposażeniu ekonomicznym średnim stanowiły 51%, natomiast w obrębie infrastruktury społecznej gorszej 49% gmin było gorzej wyposażonych w infrastrukturę ekonomiczną średnią. W obrębie kategorii infrastruktury ekonomicznej lepszej 44% stanowiły gminy średnio wyposażone w infrastrukturę techniczną, natomiast w obrębie kategorii infrastruktury ekonomicznej gorszej aż 67% stanowiły gminy gorzej wyposażone w infrastrukturę techniczną.

W tabeli 2 zawarto wybrane wyniki uzyskane z analizy korespondencji przeprowadzonej na poziomie gmin Dolnego Śląska. Między innymi zestawiono tam wartości osobliwe, wartości własne, procenty bezwładności, procenty kumulacyjne i wartości statystyki chi-kwadrat. Statystyka chi-kwadrat jest oparta na ilorazie największej wiarygodności, oceniając w ten sposób dobroć dopasowania modelu [Stanisz 2000].

Tabela 1. Rozkład profili częstości zaobserwowanych na Dolnym Śląsku (%)
 Table 1. The schedule profile of observed frequencies in the Lower Silesia area (%)

Infrastruktura Infrastructure		Techniczna Technical			Społeczna Social			Ekonomiczna Economic		
		lepsz good	średn average	gorsza worse	lepsz good	średn average	gorsza worse	lepsz good	średn average	gorsza worse
Techniczna Technical	lepsz good	100	0	0	32	35	17	31	38	8
	średn average	0	100	0	34	38	34	44	36	25
	gorsza worse	0	0	100	34	27	49	25	26	67
Społeczna Social	lepsz good	35	30	29	100	0	0	42	24	31
	średn average	43	36	24	0	100	0	36	38	25
	gorsza worse	22	34	47	0	0	100	22	38	44
Ekonomiczna Economic	lepsz good	30	34	18	37	29	17	100	0	0
	średn average	62	47	33	37	51	49	0	100	0
	gorsza worse	8	19	49	26	20	34	0	0	100

Źródło: badania własne.
 Source: own research.

W wyniku analizy korespondencji z wprowadzonych danych wyodrębnionych zostało dla Dolnego Śląska dziewięć wymiarów. Jak wynika z tabeli 2, wymiar 1 wyjaśnia tylko ponad 25% bezwładności, co oznacza, że częstości względne, które można odtworzyć z informacji jednowymiarowej, mogą odtworzyć tylko 25% ogólnej statystyki chi-kwadrat (a więc także bezwładności). Ostatecznym celem analizy korespondencji jest oczywiście znalezienie teoretycznej interpretacji znaczenia wprowadzonych wymiarów. Uzyskane dane wskazują, że cztery wymiary bardziej różnicują przypadki – gminy i pozwalają wyjaśnić ponad 77% zmienności badanej populacji Dolnego Śląska.

Uzyskane w dalszym rozwiązaniu wielowymiarowym współrzędne dla gmin Dolnego Śląska pozwoliły wyznaczyć punkty o tych współrzędnych na wielowymiarowym wykresie rozrzutu. Odległości punktów na wykresie dostarczyły informacji o podobieństwie częstości względnych, jakie dane wiersze mają w odpowiednich kolumnach. Odległości te nie są prostymi odległościami euklidesowymi obliczonymi z częstości względnych wierszy i kolumn, ale są to odległości ważone. W ocenie jakości rozwiązania o wybranej liczbie wymiarów pomocne są dodatkowe statystyki. Jedną z nich jest kwadrat cosinusa, który określa miary jakości dla każdego punktu w odniesieniu do każdego wymiaru. Suma wartości z kolumn liczona po wszystkich wymiarach jest równa

Tabela 2. Wartości własne i bezwładność dla wszystkich wymiarów w gminach Dolnego Śląska
 Table 2. Own values and inertia for all dimensions on the level of the Lower Silesia communes

Liczba wymiarów Number of dimensions	Wartości osobliwe Peculiar values	Wartości własne (bezwładność) Own values (inertia)	Procent bezwładności Percentage of inertia	Procent skumulowany Percentage accumulated	Chi-kwadrat Chi the square
Wymiar 1 Dimension 1	0,50	0,25	25,01	25,01	487,61
Wymiar 2 Dimension 2	0,45	0,20	19,83	44,84	386,67
Wymiar 3 Dimension 3	0,41	0,17	17,07	61,91	332,83
Wymiar 4 Dimension 4	0,39	0,16	15,55	77,45	303,12
Wymiar 5 Dimension 5	0,36	0,13	12,73	90,18	248,13
Wymiar 6 Dimension 6	0,31	0,10	9,82	100,00	191,55
Wymiar 7 Dimension 7	0,00	0,00	0,00	100,00	0,00
Wymiar 8 Dimension 8	0,00	0,00	0,00	100,00	0,00
Wymiar 9 Dimension 9	0,00	0,00	0,00	100,00	0,00

Źródło: badania własne.
 Source: own research.

całkowitej jakości. Wartość ta może być również interpretowana jako korelacja danego punktu z odpowiednim wymiarem. Zgodnie z przyjętą metodyką, do obliczenia sumy punktów wykorzystano współrzędne uzyskane dla czterech wymiarów, jako bardziej różnicujących analizowane przypadki (gminy). Dla uzyskanych wartości współrzędnych przyjęto wartości minimalne równe 0 i maksymalne równe 100. Ocena punktowa polegała na wyliczeniu wartości funkcji dla poszczególnych współrzędnych (sumy współrzędnych) dla czterech wymiarów i w obrębie poszczególnych kategorii jakościowych infrastruktury technicznej, społecznej i ekonomicznej.

W tabeli 3 zestawiono ocenę punktową wyposażenia infrastrukturalnego gmin Dolnego Śląska. Zawiera ona liczbę punktów ogółem wraz z podziałem na rodzaje infrastruktury. Poziom wyposażenia ogółem zawiera się w przedziale od 29 do 276 punktów. Wartość minimalną (gminy najgorzej wyposażone) osiągnęło sześć gmin, czyli 4,5% badanej populacji. Wartość maksymalną wyliczono dla siedmiu gmin, czyli dla 5,3% populacji.

Nadanie wartości punktowych również poszczególnym rodzajom infrastruktury umożliwiło ocenę jej wewnętrznego zróżnicowania. Tak więc np. gmina Twardogóra została oceniona pod względem wyposażenia infrastrukturalnego ogółem na 182 punkty.

Tabela 3. Ocena punktowa poziomu wyposażenia w infrastrukturę przykładowych gmin Dolnego Śląska

Table 3. Points estimation of the level of infrastructure equipment for exemplified Lower Silesian communities

Gminy Communities	Liczba punktów – Number of points			
	ogółem total	techniczna technological	społeczna social	ekonomiczna economical
Lądek Zdrój	276	100	84	92
Stronie Śląskie	276	100	84	92
Szczytna	228	100	36	92
Miłkowice	218	42	84	92
Kobierzyce	217	100	25	92
Twardogóra	182	100	36	46
Długołęka	170	42	36	92
Święta Katarzyna	159	42	25	92
Góra	129	100	25	4
Bolesławiec	124	42	36	46
Brzeg Dolny	82	42	36	4
Wołów	82	0	36	46
Domaniów	71	0	25	46
Jordanów Śląski	40	0	36	4
Gromadka	29	0	25	4

Źródło: badania własne.
Source: own research.

Największy udział miała infrastruktura techniczna (100 punktów) i ekonomiczna (46 punktów); relatywnie gmina ta prezentuje najniższy poziom infrastruktury społecznej (36 punktów). Zdecydowanie inna sytuacja jest w gminie Długołęka. Waloryzacja ogółem (170 punktów) nie odbiega zasadniczo od waloryzacji gminy Twardogóra, lecz infrastruktura techniczna jest tutaj gorzej rozwinięta (42 punkty), a infrastrukturę społeczną oceniono na 92 punkty.

PODSUMOWANIE I WNIOSKI

Przeprowadzone badania związane z wyposażeniem infrastrukturalnym obszarów wiejskich pozwoliły na sformułowanie następujących wniosków:

1. Zastosowana analiza skupień umożliwiła wyodrębnienie grup badanych gmin wewnątrznie jednorodnych pod względem infrastruktury technicznej, społecznej i ekonomicznej.

2. W wyniku analizy skupień utworzono kategorie jakościowe opisujące poziomy wyposażenia infrastrukturalnego gmin niezbędne do nadania punktacji.

3. Zastosowana wielowymiarowa analiza korespondencji pozwoliła na waloryzację infrastruktury obszarów wiejskich przez nadanie badanym gminom wartości liczbowych (punktów) zarówno dla całej infrastruktury, jak i z podziałem na poszczególne jej rodzaje (techniczna, społeczna i ekonomiczna), uwzględniając wcześniej wyodrębnione kategorie jakościowe.

4. Waloryzacja ta może mieć duży aspekt praktyczny, ułatwiając podejmowanie decyzji inwestycyjnych przez władze lokalne.

5. Władze samorządowe powinny szczególnie dbać o gminy o najniższej punktacji i dawać im pierwszeństwo w zamierzeniach inwestycyjnych, ze szczególnym uwzględnieniem rodzaju infrastruktury o najmniejszej liczbie punktów.

6. Ocena punktowa gmin może również pomóc podjąć decyzję przedsiębiorcom, którzy mają zamiar rozwijać działalność gospodarczą w tym rejonie.

7. Uwzględniając przyciągający charakter infrastruktury, inwestorzy prawdopodobnie chętniej zaktzywizują przedsiębiorczość na terenie gmin o lepszej waloryzacji, natomiast mniejsze zainteresowanie wykażą gminami o słabej waloryzacji w zakresie infrastruktury.

8. Posiadając informację związaną ze stanem infrastruktury, można dokonać waloryzacji opartej na zaproponowanej metodzie dla dowolnej liczby gmin, a więc można ją wykorzystać do planowania rozwoju infrastruktury dowolnego regionu kraju.

Przeprowadzona diagnoza stanu infrastruktury obszarów wiejskich Dolnego Śląska, a przede wszystkim zaproponowany sposób oceny punktowej poziomu wyposażenia w poszczególne rodzaje infrastruktury może być pomocny w realizacji polityki regionalnej. Wobec powszechnego deficytu środków finansowych na modernizację obszarów wiejskich i dużej konkurencji o ich przyznanie (także funduszy strukturalnych wspierających rozwój infrastruktury), zaproponowana metoda waloryzacji może być dobrą podstawą podejmowania decyzji związanych z alokacją środków finansowych w ramach sprawnych planów i strategii.

LITERATURA

- Kopec B., 1983. *Metodyka badań ekonomicznych w gospodarstwach rolnych*. Wyd. AR, Wrocław, 166-199.
- Pięcek B., 1997. *Infrastrukturalne uwarunkowania przedsiębiorczości na obszarach wiejskich*. W: *Ekonomiczne i społeczne uwarunkowania i możliwości wielofunkcyjnego rozwoju wsi w Polsce*. SGGW, Warszawa, 117-118.
- Program Statistica PL – operacyjny program – opisy z pomocy (analiza skupień i wielowymiarowa analiza korespondencji). 2000.
- Ryznar J., 1999. *Metody stosowane w pracy badawczej z zakresu doradztwa rolniczego*. Wyd. AR, Wrocław, 52-58.
- Sagan A., 2001. *Przykłady zaawansowanych technik analitycznych w badaniach marketingowych*. AE w Krakowie, http://www.statsoft.pl/czytelnia/marketing/spis_treści, 1-6.
- Stachak S., 2006. *Podstawy metodologii nauk ekonomicznych*. Wyd. Książka i Wiedza, Warszawa, 213-216.
- Stanisz A., 2000. *Przystępny kurs statystyki w oparciu o program Statistica PL na przykładach z medycyny*. T. 2. Statsoft, Kraków.

- Wilczyńska K., 1997. Wyposażenie infrastrukturalne wsi i rolnictwa w Polsce w latach 1989-1995 – problemy rozwoju. W: Wieś i rolnictwo w okresie przemian systemowych. AR Poznań, 103-116.
- Wilkin J., 1996. Obszary wiejskie w polityce rozwoju gospodarczego Polski. PTE, FAPA, Komitet Ekonomiki Rolnictwa, PAN, Warszawa, 32.

**APPLICATION OF CLUSTER ANALYSIS METHODS
AND MULTIDIMENSIONAL ANALYSIS CORRESPONDENCE
TO THE ESTIMATIONS LEVEL OF INFRASTRUCTURE ON RURAL AREAS**

Summary. The multifunctional development of rural areas is connected with increased number of out of agriculture functions. The level of infrastructural development plays an essential role in the process of multifunctional development of the rural areas and significantly affects the life standard of the inhabitants. The work presents the analysis of the endowment level, basic technological, social and economical infrastructure at the territory of chosen communities in the Lower Silesia area. After that on the basis of statistical methods the evaluation of points valorisation rural areas in the light of level at infrastructure has been realised with differentiation in communes.

Key words: infrastructure, rural areas, diversification, valorisation, Lower Silesia

Zaakceptowano do druku – Accepted for print: 18.04.2009

Do cytowania – For citation: Kropisz I., 2009. Zastosowanie metody analizy skupień oraz wielowymiarowej analizy korespondencji do oceny poziomu infrastruktury obszarów wiejskich. J. Agribus. Rural Dev. 3(13), 129-137.