

STATYSTYKA W PRAKTYCE

Andrzej BAŁ

Zastosowanie metod wielowymiarowej analizy porównawczej do oceny stanu środowiska w województwie dolnośląskim¹

Streszczenie. *Celem opracowania jest zaprezentowanie wyników metod wielowymiarowej analizy porównawczej wykorzystywanej do oceny stanu środowiska w województwie dolnośląskim w przekroju powiatów. Badanie przeprowadzono na podstawie danych GUS za 2015 r. dotyczących stanu i ochrony środowiska w 30 powiatach województwa dolnośląskiego. W badaniu zastosowano metody porządkowania liniowego obiektów (wzorcowe i bezwzorcowe). Wiele z nich opisanych w literaturze przedmiotu na ogół prowadzi do zróżnicowanych wyników (rankingi obiektów nie są takie same). Wynika to m.in. z przyjętych metod normalizacji i ważenia zmiennych oraz agregacji (tworzenia zmiennych syntetycznych). W artykule podjęto próbę porównania wyników porządkowania liniowego powiatów ze względu na stan środowiska za pomocą wybranych metod wzorcowych i bezwzorcowych. W analizie poprawności rankingów wykorzystano mierniki oceny jakości metod porządkowania liniowego.*

Słowa kluczowe: ochrona środowiska, porządkowanie liniowe, metoda Helwiga, metoda TOPSIS, główne składowe, program R.

JEL: Q00, Q59, C19, C88

Zanieczyszczenie środowiska jest skutkiem działalności człowieka, m.in. przemysłowej i rolniczej. Intensywnie zużywane są nieodnawialne zasoby naturalne, niszczone są lasy, ginie wiele gatunków zwierząt. Skażeniu ulegają

¹ Artykuł opracowano na podstawie referatu wygłoszonego na konferencji pt. *Rozwój gospodarczy i przestrzenny Polski a realizacja polityki spójności*, która odbyła się 25 i 26 maja 2017 r. w Katowicach.

gleba, woda i powietrze. Wszystkie kontynenty — poza Antarktydą — są w znacznym stopniu zanieczyszczone. Szacuje się, że jedynie od kilku do kilkudziesięciu procent obszaru poszczególnych z nich pozostało niezmiennych wskutek działalności człowieka (Europa — ok. 3%, Azja — ok. 14%, Ameryka Południowa — ok. 21%, Afryka — ok. 27%, Australia — ok. 28%, Ameryka Północna — ok. 38%).

Polska należy do grupy państw o bardzo dużym zanieczyszczeniu środowiska, istnieje zatem potrzeba monitorowania jego stanu i działań zapobiegających temu procesowi. Tej problematyce w naszym kraju poświęcono liczne ustawy, m.in. o ochronie środowiska, o ochronie przyrody, o ochronie zwierząt, o odpadach oraz prawo wodne. Ta tematyka jest obecna również w konstytucji.

Zanieczyszczenie środowiska w regionach Polski jest zróżnicowane. W artykule przedstawiono wyniki badania stanu środowiska w powiatach województwa dolnośląskiego w 2015 r. z zastosowaniem wielowymiarowej analizy porównawczej, do której zalicza się metody porządkowania liniowego, wykorzystywane w badaniach ekonomicznych w celu ustalenia kolejności (rankingu) lub klasyfikacji obiektów wielocechowych, takich jak kraje lub regiony (ze względu na poziom rozwoju gospodarczego), przedsiębiorstwa (ze względu na kondycję finansową) czy produkty (ze względu na walory użytkowe).

Idea porządkowania liniowego obiektów wielowymiarowych opiera się na pojęciu porządkującej relacji binarnej (zwrotnej, antysymetrycznej, przechodniej i spójnej). Z aksjomatów tej relacji wynika, że jest możliwe stwierdzenie, który z dwóch dowolnych obiektów zbioru jest pierwszy (lepszy), a który drugi (gorszy), a także czy są one identyczne. Przedmiotem porządkowania liniowego są obiekty lub zjawiska opisane przez wiele zmiennych, których wartości są mierzone i gromadzone w zbiorach danych statystycznych. Takie cechy obiektów i zjawisk ekonomicznych, jak rozwój gospodarczy, kondycja finansowa oraz walory użytkowe produktów lub usług są natomiast zmiennymi, których wartości nie można bezpośrednio zmierzyć. Charakterystyka tych zmiennych opiera się na funkcjach obserwacji bezpośrednio mierzalnych cech diagnostycznych (funkcje agregujące mogą mieć różną postać analityczną). Uzyskane wyniki zmiennej syntetycznej umożliwiają uporządkowanie obiektów wielowymiarowych według preferencji (dominacji).

Celem artykułu jest ocena stanu środowiska w województwie dolnośląskim w przekroju powiatów w 2015 r. na podstawie publikowanych danych statystycznych GUS z wykorzystaniem wybranych metod wielowymiarowej analizy porównawczej. W artykule przedstawiono także historycznie pierwsze metody porządkowania liniowego opracowane w różnych obszarach badawczych oraz porównano wyniki porządkowania obiektów za pomocą tych procedur.

W obliczeniach i wizualizacji wyników badań posłużono się programem CRAN R oraz niektórymi pakietami programu R. W analizach porównawczych metod porządkowania liniowego wykorzystano pakiety programu R — *pllord* (Bąk, 2015) i *topsis* (Yazdi, 2015) oraz opracowane skrypty w języku R. Program R (R Development Core Team, 2017) jest niekomercyjnym projektem analizy danych, powszechnie wykorzystywanym m.in. w statystycznych i ekonometrycznych badaniach ekonomicznych.

METODY PORZĄDKOWANIA LINIOWEGO

Wielowymiarowa analiza porównawcza umożliwia ocenę obiektów i zjawisk złożonych, czyli takich, na których stan i zachowanie wpływa jednocześnie wiele cech (zmiennych) i czynników. Zwięzła definicja podana przez Hellwiga mówi, że (...) *metody i technika porównywania obiektów wielocechowych nazywają się wielowymiarową analizą porównawczą* (...) (Hellwig, 1981, s. 48).

Podstawowym celem wielowymiarowej analizy porównawczej jest konstrukcja miary syntetycznej umożliwiającej porównywanie elementów zbioru (obiektów) opisanych za pomocą wielu zmiennych (cech). Aby go osiągnąć, często wykorzystuje się metody porządkowania liniowego. W metodach tych zakłada się, że rezultatem uzyskanym na skutek zastosowania odpowiedniego algorytmu będzie takie uszeregowanie zbioru obiektów, w którym (Grabiński, 1992, s. 135):

- każdy obiekt ma przynajmniej jednego sąsiada oraz nie więcej niż dwóch sąsiadów,
- jeżeli obiekt *a* jest sąsiadem obiektu *b*, to obiekt *b* jest sąsiadem obiektu *a*,
- są tylko dwa obiekty mające jednego sąsiada.

Metody porządkowania liniowego prowadzą do ustalenia kolejności obiektów ze względu na jedną agregatową cechę, która jest syntetycznym reprezentantem wielu cech opisujących porządkowane obiekty. W ujęciu geometrycznym metody porządkowania liniowego prowadzą do rzutowania punktów reprezentujących obiekty umieszczone w wielowymiarowej przestrzeni cech na prostą, która reprezentuje cechę agregatową.

Podstawowe kategorie w porządkowaniu liniowym to obiekty i cechy. Zarówno obiekty, jak i cechy posiadają swoje liczbowe reprezentacje (obrazy), co umożliwia ich wszechstronne analizowanie za pomocą algorytmów numerycznych. Porządkowanie liniowe ma charakter wartościujący (oceniający) elementy (obiekty) i znajduje zastosowanie w różnych dziedzinach badań, w tym w ekonomii.

Metody porządkowania liniowego, mieszczące się w obrębie wielowymiarowej analizy porównawczej i szerzej taksonomii², są w dużej mierze dorobkiem polskiej myśli statystycznej i ekonometrycznej. Pierwszą propozycję porządkowania liniowego przedstawił Hellwig (1968). Publikacja ta zainicjowała intensywne badania, których efektem były kolejne propozycje metod porządkowania liniowego (Bartosiewicz, 1976; Borys, 1978b; Cieślak, 1974; Nowak, 1984; Pluta, 1976; Strahl, 1978; Walesiak, 1993).

Podstawą porządkowania liniowego jest zmienna syntetyczna³, której wartości są szacowane na podstawie obserwacji zmiennych diagnostycznych opisujących badane obiekty. Zmienna syntetyczna ma charakter zmiennej ukrytej, ponieważ jej wartości nie są obserwowane bezpośrednio. Jest to natomiast możliwe dzięki obserwacji zmiennych diagnostycznych, które są bezpośrednio mierzalne. Wartości zmiennej syntetycznej szacuje się za pomocą funkcji agregującej

² Taksonomia, wielowymiarowa analiza porównawcza i pojęcia bliskoznaczne zostały szerzej omówione w literaturze (Pociecha, Podolec, Sokołowski i Zajac, 1988; Pociecha, 2008).

³ W literaturze przedmiotu można także spotkać inne określenia zmiennej syntetycznej, np. zmienna agregatowa, miara syntetyczna, syntetyczna miara rozwoju, taksonomiczny miernik rozwoju, agregatowa miara rozwoju, miara rozwoju gospodarczego czy metacecha.

cych, których postać analityczna może być różna. Rozróżnia się dwie podstawowe grupy metod wykorzystywanych do szacowania wartości zmiennej syntetycznej — bezwzorcowe i wzorcowe.

W procedurze porządkowania liniowego wyróżnia się takie etapy postępowania, jak: określenie charakteru zmiennych (stymulanty, nominanty, zdestymulanty)⁴, wyznaczanie wag zmiennych i ich normalizacja, wyznaczenie współrzędnych wzorca w przypadku agregacji wzorcowej, agregacja bezwzorcowa lub wzorcowa, klasyfikacja porangowanych obiektów i rozpoznanie typów rozwojowych (Bąk, 1999, 2013; Grabiński, 1984; Wysocki, 2010, s. 145).

Pierwsza propozycja zastosowania metody porządkowania liniowego — umożliwiającej porządkowanie liniowe obiektów z wykorzystaniem wzorca — w obszarze badań taksonomicznych i ekonomicznych została przedstawiona przez Hellwiga w 1968 r. pod nazwą miara rozwoju gospodarczego (Hellwig, 1968). Rok wcześniej (w 1967 r.) taka koncepcja została opisana w niepublikowanym raporcie UNESCO zatytułowanym *Procedure of Evaluating High Level Manpower Data and Typology of Countries by Means of the Taxonomic Method*.

Na gruncie teorii decyzji (wielokrotnego podejmowania decyzji) pierwszą metodę porządkowania liniowego z wykorzystaniem wzorca i antywzorca zaproponowali Hwang i Yoon (1981) pod nazwą TOPSIS — *Technique for Order Preference by Similarity to Ideal Solution*, a opisali ją rok wcześniej (Yoon, 1980; Yoon i Hwang, 1980).

Ideę zastosowania analizy głównych składowych (PCA — *Principal Component Analysis*) do porządkowania liniowego obiektów (na podstawie wartości pierwszej głównej składowej) zaproponował w badaniach przyrodniczych i rolniczych Perkal (1967), a jej podstawy zostały już przedstawione w pracach Pearsona (1901) i Hotellinga (1933).

Porządkowanie liniowe według Hellwiga i TOPSIS to metody wzorcowe, przy czym w pierwszej z nich punktem odniesienia obiektów w przestrzeni wielowymiarowej jest wzorzec, a w drugiej wyznaczane są dwa punkty odniesienia — wzorzec i antywzorzec. Metoda wykorzystująca wartości pierwszej głównej składowej jest metodą bezwzorcową, opierającą się na wartościach i wektorach własnych macierzy kowariancji lub korelacji.

Konstrukcja miary syntetycznej (miary rozwoju gospodarczego) Hellwiga przedstawia się następująco (Hellwig, 1968):

a) normalizacja zmiennych (standaryzacja): $z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$, x_{ij} — obserwacja j -tej zmiennej dla obiektu i , \bar{x}_j — średnia arytmetyczna obserwacji j -tej zmiennej, s_j — odchylenie standardowe obserwacji j -tej zmiennej;

b) współrzędne wzorca:

$$z_{0j} = \begin{cases} \max_i \{z_{ij}\} & \text{dla zmiennych o charakterze stymulant} \\ \min_i \{z_{ij}\} & \text{dla zmiennych o charakterze destymulant;} \end{cases}$$

c) odległości obiektów od wzorca: $d_{i0} = \sqrt{\sum_{j=1}^m (z_{ij} - z_{0j})^2}$;

⁴ Pojęcia zmiennych o charakterze stymulant i destymulant zostały wprowadzone do literatury przez Hellwiga (1968), a pojęcie zmiennej o charakterze nominanty — przez Borysa (1978a).

- d) wartości zmiennej agregatowej: $q_i = 1 - \frac{d_{i0}}{d_0}$, przy czym: na ogół $q_i \in [0; 1]$ ⁵,
 $\max_i\{q_i\}$ — najlepszy obiekt, $\min_i\{q_i\}$ — najgorszy obiekt, $d_0 = \bar{d}_0 + 2s_d$, $\bar{d}_0 =$
 $= \frac{1}{n} \sum_{i=1}^n d_{i0}$, $s_d = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_{i0} - \bar{d}_0)^2}$.

Konstrukcja TOPSIS Hwanga i Yoona (Hwang i Yoon, 1981; Yoon i Hwang, 1995) to:

- a) normalizacja zmiennych (przekształcenie ilorazowe): $z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}}$, x_{ij} — ob-

serwacja j -tej zmiennej dla obiektu i ;

- b) współrzędne wzorca:

$$z_{0j}^+ = \begin{cases} \max_i\{z_{ij}\} & \text{dla zmiennych stymulant} \\ \min_i\{z_{ij}\} & \text{dla zmiennych destymulant;} \end{cases}$$

- c) współrzędne antywzorca:

$$z_{0j}^- = \begin{cases} \min_i\{z_{ij}\} & \text{dla zmiennych stymulant} \\ \max_i\{z_{ij}\} & \text{dla zmiennych destymulant;} \end{cases}$$

- d) odległości obiektów od wzorca: $d_{i0}^+ = \sqrt{\sum_{j=1}^m (z_{ij} - z_{0j}^+)^2}$;

- e) odległości obiektów od antywzorca: $d_{i0}^- = \sqrt{\sum_{j=1}^m (z_{ij} - z_{0j}^-)^2}$;

- f) wartości zmiennej agregatowej: $q_i = \frac{d_{i0}^-}{d_{i0}^+ + d_{i0}^-}$, przy czym: $q_i \in [0; 1]$, $\max_i\{q_i\}$ — najlepszy obiekt, $\min_i\{q_i\}$ — najgorszy obiekt.

Konstrukcja PCA jest zaś następująca (Perkal, 1967; Rusnak, 1999; Balicki, 2009):

- a) normalizacja zmiennych (standaryzacja): $z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$, x_{ij} — obserwacja j -tej zmiennej dla obiektu i , \bar{x}_j — średnia arytmetyczna obserwacji j -tej zmiennej, s_j — odchylenie standardowe obserwacji j -tej zmiennej lub wykorzystanie danych pierwotnych (bez normalizacji);
- b) obliczenie macierzy kowariancji **S** lub korelacji **R**;
- c) obliczenie wartości własnych i wektorów własnych macierzy kowariancji **S** lub korelacji **R**;
- d) zmiana znaków współczynników pierwszej głównej składowej (elementów pierwszego wektora własnego) — jeżeli więcej jest współczynników ujemnych niż dodatnich⁶;
- e) obliczenie wartości składowych głównych na podstawie danych empirycznych i wektorów własnych: $\mathbf{Y} = \mathbf{XW}$: **Y** — macierz głównych składowych,

⁵ Znormalizowane wartości zmiennej agregatowej Hellwiga q_i mogą wykroczyć poza przedział $[0; 1]$, jeżeli w zbiorze porządkowanych obiektów występują obiekty istotnie odstające od innych (wartości zmiennej agregatowej spoza przedziału $[0; 1]$ będą wskazywać na te obiekty).

⁶ Zmiana znaków współczynników pierwszej głównej składowej wpływa na kierunek porządkowania obiektów — tutaj większe wartości pierwszej głównej składowej oznaczają wyższą pozycję powiatu.

- X — macierz danych empirycznych, W — macierz współczynników głównych składowych (wektory własne);
- f) uporządkowanie malejąco obiektów na podstawie wartości pierwszej składowej głównej.

BADANIE EMPIRYCZNE

Badanie przeprowadzono na podstawie danych statystycznych o stanie i ochronie środowiska w 30 powiatach województwa dolnośląskiego w 2015 r., udostępnionych na stronach internetowych Urzędu Statystycznego we Wrocławiu. Uwzględniono wstępnie 10 zmiennych, które pokazano w zestawieniu (1).

**ZESTAWIENIE (1) ZMIENNYCH OBRAZUJĄCYCH STAN I OCHRONĘ ŚRODOWISKA
W WOJEWÓDZTWIE DOLNOŚLĄSKIM W 2015 R.**

Zmienne	Opis	Charakter zmiennej
X_1	pobór wody na potrzeby gospodarki narodowej i ludności	destymulanta
X_2	zużycie wody na potrzeby gospodarki narodowej i ludności	destymulanta
X_3	ścieki przemysłowe i komunalne oraz ludność korzystająca z oczyszczalni ścieków	destymulanta
X_4	oczyszczalnie ścieków	stymulanta
X_5	emisja i redukcja zanieczyszczeń powietrza	destymulanta
X_6	ochrona przyrody i krajobrazu — powierzchnia	stymulanta
X_7	ochrona przyrody i krajobrazu — pomniki przyrody	stymulanta
X_8	odpady wytworzone i dotychczas składowane oraz tereny ich składowania	destymulanta
X_9	nakłady na środki trwałe służące ochronie środowiska	stymulanta
X_{10}	nakłady na środki trwałe służące gospodarce wodnej	stymulanta

Ź r ó ł o: opracowanie własne na podstawie danych Urzędu Statystycznego we Wrocławiu.

W badaniu zastosowano procedurę obejmującą:

- 1) określenie charakteru zmiennych (stymulanty, nominanty, destymulanty). Na podstawie analizy merytorycznej wskazano charakter zmiennych, wśród których znalazły się stymulanty i destymulanty — zestawienie (1). Cech o charakterze nominant nie zidentyfikowano. Destymulanty przekształcono do postaci zmiennych stymulant z zastosowaniem równania: $x_{ij}^S = 1 - x_{ij}^D$;
- 2) formalny dobór zmiennych. Zmienne opisujące objekty w porządkowaniu liniowym powinny spełniać następujące warunki: dobrze dyskryminować objekty (charakteryzować się dużą zmiennością), być słabo skorelowane między sobą oraz być mocno skorelowane ze zmiennymi odrzuconymi. W celu doboru optymalnego podzbioru zmiennych zastosowano metodę odwróconej macierzy korelacji Maliny i Zeliasia (1997). W wyniku analizy odrzucono cztery zmienne — X_1 , X_2 , X_5 , X_8 i w dalszych badaniach uwzględniono sześć — X_3 , X_4 , X_6 , X_7 , X_9 i X_{10} ;

- 3) ustalenie wag zmiennych — przyjęto takie same (jednostkowe) wagi dla wszystkich zmiennych (Zeliaś, 2000);
- 4) normalizację danych metodą standaryzacji dla wszystkich wykorzystanych metod;
- 5) wyznaczenie współrzędnych wzorca — w przypadku metody Hellwiga, a w przypadku metody TOPSIS — wyznaczenie współrzędnych wzorca i antywzorca na podstawie macierzy danych standaryzowanych;
- 6) obliczenie współczynników głównych składowych na podstawie macierzy korelacji — w przypadku metody PCA (Rusnak, 1999; Balicki, 2009);
- 7) porządkowanie liniowe badanych obiektów (powiatów) z zastosowaniem metod Hellwiga, TOPSIS i PCA;
- 8) ocenę jakości rankingów badanych powiatów na podstawie mierników jakości procedur porządkowania liniowego⁷ (Grabiński, Wydymus i Zeliaś, 1989; Bąk, 1999).

Wyniki porządkowania liniowego powiatów województwa dolnośląskiego ze względu na stan ochrony środowiska w 2015 r. zaprezentowano w tabl. 1 i 2. W tabl. 3 przedstawiono ocenę zgodności uporządkowania (rankingów) metodami Hellwiga, TOPSIS i PCA mierzoną współczynnikiem korelacji rang Spearmana.

**TABL. 1. WYNIKI PORZĄDKOWANIA LINIOWEGO POWIATÓW
WOJEWÓDZTWA DOLNOŚLĄSKIEGO**

Powiaty	Wartości miary syntetycznej na podstawie metody		
	Hellwiga	TOPSIS	PCA
Bolesławiecki	0,1550	0,3419	-0,2626
Dzierżoniowski	0,1480	0,3485	-0,3434
Głogowski	0,1167	0,3204	0,6483
Górowski	0,1454	0,4014	-1,6515
Jaworski	0,1341	0,3563	-0,8789
Jeleniogórski	0,1628	0,3601	-0,2237
Kamiennogórski	0,0931	0,3396	-0,4721
Kłodzki	0,2626	0,4365	-0,7045
Legnicki	0,1316	0,3568	-0,7420
Lubański	0,1106	0,3518	-0,6206
Lubiński	0,0877	0,2436	1,8260
Lwówecki	0,0917	0,3391	-0,7198
Milicki	0,1453	0,3972	-1,3675
Oleśnicki	0,1411	0,3478	-0,2256
Oławski	0,0682	0,3314	-0,3063
Polkowicki	0,1501	0,3631	-0,4316
Strzeliński	0,0703	0,3372	-0,7670
Średzki	0,1144	0,3548	-0,1143
Świdnicki	0,1651	0,3439	0,0861
Trzebnicki	0,2650	0,4171	-0,2243

⁷ Mierniki jakości dotyczą pomiaru i oceny takich własności procedur porządkowania liniowego, jak: zgodność odwzorowania, korelacja liniowa pomiędzy zmienną syntetyczną a zmiennymi diagnostycznymi, korelacja rangowa zmiennej syntetycznej ze zmiennymi diagnostycznymi, zmienność i koncentracja zmiennej syntetycznej oraz przeciętna odległość taksonomiczna zmiennej syntetycznej od zmiennych diagnostycznych. Mierniki cząstkowe zdefiniowano tak, aby ich kierunek preferencji był jednakowy. Mniejsze wartości mierników wskazują na lepszą procedurę porządkowania liniowego.

**TABL. 1. WYNIKI PORZĄDKOWANIA LINIOWEGO POWIATÓW
WOJEWÓDZTWA DOLNOŚLĄSKIEGO (dok.)**

Powiaty	Wartości miary syntetycznej na podstawie metody		
	Hellwiga	TOPSIS	PCA
Wałbrzyski	0,1598	0,3688	-0,7537
Wołowski	0,0936	0,3155	-0,1982
Wrocławski	0,2109	0,3593	0,4492
Ząbkowicki	0,1527	0,3619	-0,4214
Zgorzelecki	0,1907	0,4104	3,7351
Złotoryjski	0,0986	0,3434	-0,6303
Jelenia Góra	0,0222	0,2984	-0,4171
Legnica	0,0525	0,3057	-0,3602
Wrocław	0,3244	0,5643	6,9494
Wałbrzych	0,0161	0,3108	-0,8573

Ź r ó d ł o: jak przy zestawieniu (1).

**TABL. 2. RANKING POWIATÓW WOJEWÓDZTWA DOLNOŚLĄSKIEGO
NA PODSTAWIE METOD**

Pozycja w rankingu	Hellwiga	TOPSIS	PCA
1	Wrocław	Wrocław	Wrocław
2	trzebnicki	kłodzki	zgorzelecki
3	kłodzki	kłodzki	lubiński
4	wrocławski	zgorzelecki	głogowski
5	zgorzelecki	głogowski	wrocławski
6	świdnicki	milicki	świdnicki
7	jeleniogórski	wałbrzyski	średzki
8	wałbrzyski	polkowicki	wołowski
9	bolesławiecki	ząbkowicki	jeleniogórski
10	ząbkowicki	jeleniogórski	trzebnicki
11	polkowicki	wrocławski	oleśnicki
12	dzierżoniowski	legnicki	bolesławiecki
13	górowski	górowski	oławski
14	milicki	średzki	dzierżoniowski
15	oleśnicki	lubański	Legnica
16	jaworski	dzierżoniowski	Jelenia Góra
17	legnicki	oleśnicki	ząbkowicki
18	głogowski	świdnicki	polkowicki
19	średzki	złotoryjski	kamiennogórski
20	lubański	bolesławiecki	lubański
21	złotoryjski	kamiennogórski	złotoryjski
22	wołowski	lwówecki	kłodzki
23	kamiennogórski	strzeliński	lwówecki
24	lwówecki	oławski	legnicki
25	lubiński	głogowski	wałbrzyski
26	strzeliński	wołowski	strzeliński
27	oławski	Wałbrzych	Wałbrzych
28	Legnica	Legnica	jaworski
29	Jelenia Góra	Jelenia Góra	milicki
30	Wałbrzych	lubiński	górowski

Ź r ó d ł o: jak przy zestawieniu (1).

Uzyskane wyniki charakteryzują się istotnym zróżnicowaniem, przy czym większa zbieżność dotyczy rankingów otrzymanych za pomocą metod wzorcowych (Hellwiga i TOPSIS) niż między rankingami uzyskanymi za pomocą tych

metod a rankingiem otrzymanym z zastosowaniem bezwzorcowej metody porządkowania liniowego opartej na pierwszej głównej składowej (PCA). Jedynie powiat Wrocław znalazł się na tym samym (pierwszym) miejscu w przypadku wszystkich zastosowanych metod.

TABL. 3. WARTOŚCI WSPÓŁCZYNNIKÓW KORELACJI RANG SPEARMANA WEDŁUG POSZCZEGÓLNYCH METOD

Metody	Hellwiga	TOPSIS	PCA
Hellwiga	1,000	0,179	0,105
TOPSIS	0,179	1,000	-0,196
PCA	0,105	-0,196	1,000

Źródło: jak przy zestawieniu (1).

W zestawieniu (2) przedstawiono wyniki klasyfikacji powiatów województwa dolnośląskiego otrzymane na podstawie miar syntetycznych Hellwiga, TOPSIS i PCA. Wyodrębniono cztery klasy liczące odpowiednio 8, 7, 7 i 8 powiatów. W klasie pierwszej znalazły się powiaty wyróżniające się najlepszym stanem ochrony środowiska (zgodna klasyfikacja w przypadku powiatów Wrocław i zgorzeleckiego), a w czwartej — najgorszym (zgodna klasyfikacja w przypadku powiatów strzelińskiego i Wałbrzych) w świetle uwzględnionych w badaniu zmiennych i metod porządkowania liniowego.

ZESTAWIENIE (2) KLASYFIKACJI POWIATÓW NA PODSTAWIE MIAR SYNTETYCZNYCH

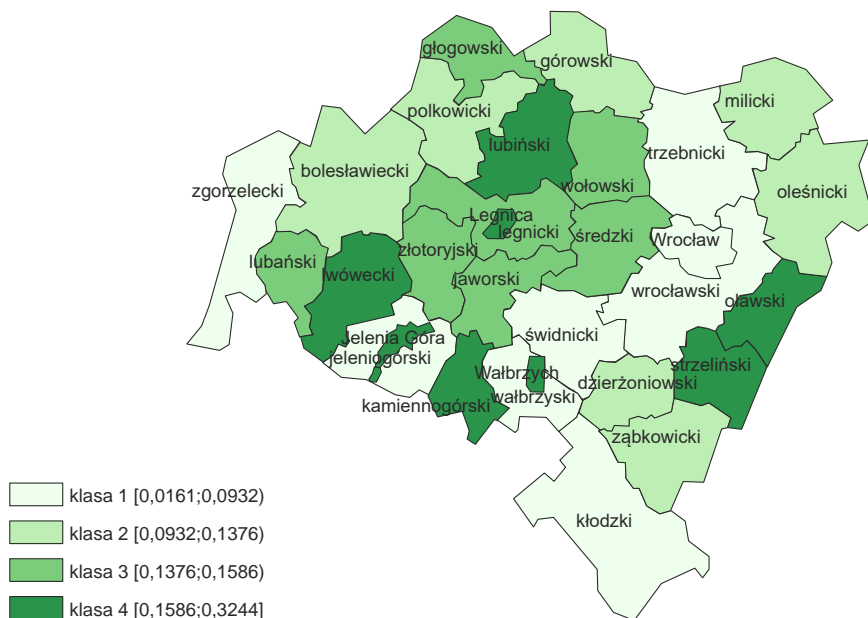
Klasa	Hellwiga	TOPSIS	PCA
I	Wrocław, trzebnicki, kłodzki, wrocławski, zgorzelecki, świdnicki, jeleniogórski, wałbrzyski	Wrocław, kłodzki, trzebnicki, zgorzelecki, górowski, milicki, wałbrzyski, polkowicki	Wrocław, zgorzelecki, lubiński, głogowski, wrocławski, świdnicki, średzki, wołowski
II	bolesławiecki, ząbkowicki, polkowicki, dzierżoniowski, górowski, milicki, oleśnicki	ząbkowicki, jeleniogórski, wrocławski, legnicki, jaworski, średzki, lubański	jeleniogórski, trzebnicki, oleśnicki, bolesławiecki, olawski, dzierżoniowski, Legnica
III	jaworski, legnicki, głogowski, średzki, lubański, złotoryjski, wołowski	dzierżoniowski, oleśnicki, świdnicki, złotoryjski, bolesławiecki, kamiennogórski, lwówecki	Jelenia Góra, ząbkowicki, polkowicki, kamiennogórski, lubański, złotoryjski, kłodzki
IV	kamiennogórski, lwówecki, lubiński, strzeliński, olawski, Legnica, Jelenia Góra, Wałbrzych	strzeliński, olawski, głogowski, wołowski, Wałbrzych, Legnica, Jelenia Góra, lubiński	lwówecki, legnicki, wałbrzyski, strzeliński, Wałbrzych, jaworski, milicki, górowski

Źródło: jak przy zestawieniu (1).

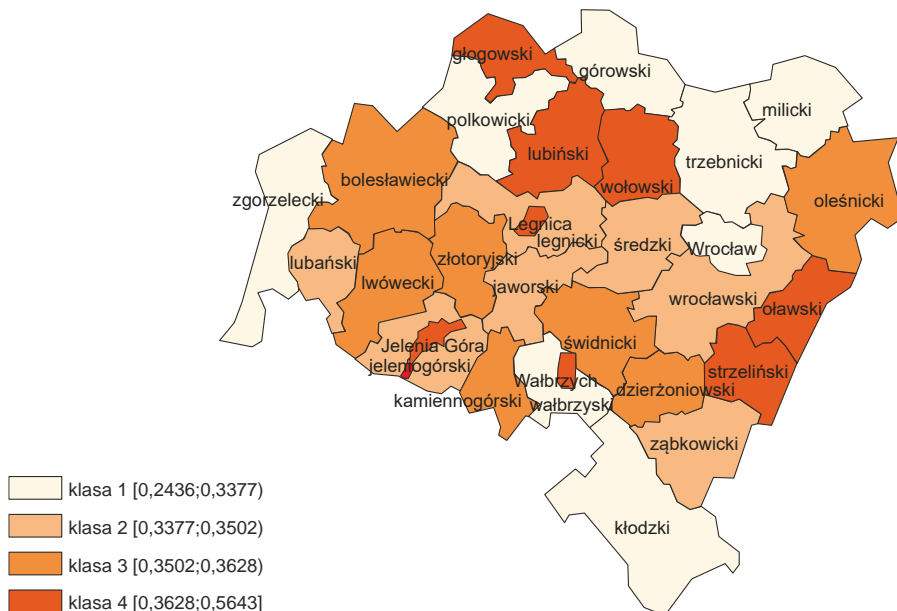
Na mapach 1—3 zilustrowano klasyfikację powiatów ze względu na stan ochrony środowiska (mniej intensywny kolor oznacza klasę powiatów charakteryzujących się lepszym stanem ochrony środowiska). Klasyfikację powiatów przeprowadzono na podstawie kwartyli, tzn. granice pierwszej klasy wyznaczono przez minimum oraz pierwszy kwartyl, drugiej — przez pierwszy kwartyl i medianę, trzeciej — przez medianę i trzeci kwartyl i wreszcie czwartej — przez trzeci kwartyl i maksimum (w legendzie map podano odpowiednie przedziały wartości zmiennych syntetycznych w poszczególnych klasach).

STAN OCHRONY ŚRODOWISKA W POWIATACH WOJEWÓDZTWA DOLNOŚLĄSKIEGO W 2015 R.

Według metody Hellwiga

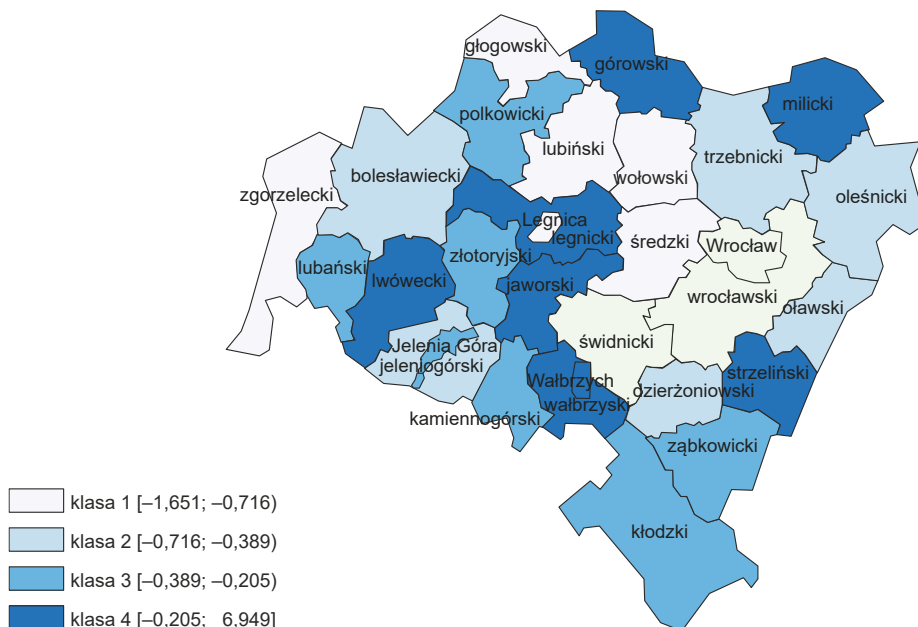


Według metody Topsis



STAN OCHRONY ŚRODOWISKA W POWIATACH WOJEWÓDZTWA DOLNOŚLĄSKIEGO W 2015 R. (dok.)

Według metody PCA



Źródło: opracowanie własne na podstawie danych US we Wrocławiu.

W ocenie jakości rankingów uwzględniono następujące mierniki oceny jakości procedur porządkowania liniowego: zgodności odwzorowania, korelacji liniowej zmiennej syntetycznej ze zmiennymi diagnostycznymi, korelacji rangowej zmiennej syntetycznej ze zmiennymi diagnostycznymi oraz zmienności i koncentracji zmiennej syntetycznej. Analityczne formuły tych mierników zamieszczają w swoich pracach Grabiński (1984), Grabiński, Wydymus i Zeliaś (1989) oraz Bąk (1999).

Mierniki oceny jakości procedur porządkowania liniowego (Hellwiga, TOPSIS i PCA) mają charakter cząstkowy o jednoznacznym kierunku preferencji — mniejsze wartości liczbowe każdego miernika wskazują na lepszą procedurę porządkowania liniowego. Można zatem przeprowadzić agregację mierników cząstkowych na podstawie wzoru (Seidler, Badach i Molisz, 1980):

$$G = \sqrt{\sum_{l=1}^q g_l^2}$$

gdzie:

G — miernik agregatowy dla metod porządkowania liniowego,

g_l — miernik cząstkowy ($l = 1, \dots, g$),

q — liczba mierników cząstkowych.

W tabl. 4 przedstawiono wartości agregatowego miernika jakości procedur porządkowania liniowego wykorzystanych w badaniu. Metoda bezwzorcowa PCA ma najniższą wartość miernika agregatowego, a więc w świetle zastosowanych cząstkowych kryteriów oceny jakości powoduje najmniejsze straty informacji w redukcji wymiarowości.

**TABL. 4. OCENA JAKOŚCI METOD
PORZĄDKOWANIA LINIOWEGO**

Metody	Agregatowy miernik jakości
Hellwiga	27,5
TOPSIS	28,1
PCA	14,8

Podsumowanie

Metody porządkowania liniowego stosuje się często w badaniach ekonomicznych, aby porównywać i oceniać obiekty i zjawiska wielocechowe (produkty, usługi, postawy czy preferencje). W artykule przedstawiono możliwości wykorzystania tych metod do oceny stanu środowiska w powiatach województwa dolnośląskiego na podstawie danych statystycznych. Wyniki badań z zastosowaniem trzech metod porządkowania liniowego (Hellwiga, TOPSIS, PCA) nie są jednoznaczne, przy czym podobieństwo rankingów mierzone współczynnikiem korelacji rang Spearmana jest większe w przypadku metod wzorcowych Hellwiga i TOPSIS.

W przypadku zastosowanych metod na pierwszej pozycji rankingowej w przeprowadzonym badaniu uplasował się Wrocław — miasto na prawach powiatu (najlepszy stan ochrony środowiska).

W wyniku klasyfikacji przeprowadzonej na podstawie wartości miar syntetycznych wyodrębniono cztery klasy. W pierwszej znalazły się powiaty charakteryzujące się najlepszym stanem ochrony środowiska (zgodna klasyfikacja w przypadku powiatów Wrocław i zgorzeleckiego), natomiast w czwartej — najgorszym (zgodna klasyfikacja w przypadku powiatów strzelińskiego i miasta Wałbrzych).

W świetle mierników jakości procedur porządkowania liniowego najlepsza okazała się bezwzorcowa metoda PCA, a następnie metody wzorcowe Hellwiga i TOPSIS.

Wyniki wskazują na kierunki nowych badań, do których można zaliczyć m.in.: porównanie wyników uporządkowania na podstawie większej liczby zbiorów danych empirycznych i większej liczby zmiennych opisujących stan ochrony środowiska, porównanie wyników porządkowania liniowego na podstawie zbiorów danych symulacyjnych o określonych rozkładach statystycznych, przeprowadzenie analizy w ujęciu dynamicznym w określonym przedziale czasu w celu poznania tendencji zmian oraz porównanie wyników rankingów na podstawie innych mierników jakości metod porządkowania liniowego.

LITERATURA

- Balicki, A. (2009). *Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne*. Wydawnictwo Uniwersytetu Gdańskiego.
- Bartosiewicz, S. (1976). Propozycja metody tworzenia zmiennych syntetycznych. *Prace Naukowe Akademii Ekonomicznej we Wrocławiu*, 84.
- Bąk, A. (1999). *Modelowanie symulacyjne wybranych algorytmów wielowymiarowej analizy porównawczej w języku C++*. Wydawnictwo Akademii Ekonomicznej we Wrocławiu.
- Bąk, A. (2013). Metody porządkowania liniowego w polskiej taksonomii — pakiet pllord. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 278, 54—62.
- Bąk, A. (2015). *Linear ordering methods — package pllord*. URL: <http://keii.ue.wroc.pl/pllord>.
- Borys, T. (1978a). Metody normowania cech w statystycznych badaniach porównawczych. *Przegląd Statystyczny*, 2, 227—239.
- Borys, T. (1978b). Propozycja agregatowej miary rozwoju obiektów. *Przegląd Statystyczny*, 3, 371—381.
- Cieślak, M. (1974). Taksonomiczna procedura prognozowania rozwoju gospodarczego i określania potrzeb na kadry kwalifikowane. *Przegląd Statystyczny*, 1, 29—39.
- Grabiński, T. (1984). Wielowymiarowa analiza porównawcza w badaniach dynamiki zjawisk ekonomicznych. *Zeszyty Naukowe Akademii Ekonomicznej w Krakowie*. Seria specjalna: *Monografie*, 61.
- Grabiński, T. (1992). *Metody taksonometrii*. Wydawnictwo Akademii Ekonomicznej w Krakowie.
- Grabiński, T., Wydymus, S., Zeliaś, A. (1989). *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*. Warszawa: PWN.
- Hellwig, Z. (1968). Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr. *Przegląd Statystyczny*, 4, 307—327.
- Hellwig, Z. (1981). Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych. W: W. Welfe (red.), *Metody i modele ekonomiczno-matematyczne w doskonaleniu zarządzania gospodarką socjalistyczną*. Warszawa: PWE.
- Hottelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(7), 417—441, 498—520.
- Hwang, C. L., Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications*. Springer-Verlag: Nowy Jork.
- Malina, A., Zeliaś, A. (1997). O budowie taksonomicznej miary jakości życia. *Taksonomia*, 4.
- Nowak, E. (1984). *Problemy doboru zmiennych do modelu ekonometrycznego*. Warszawa: PWN.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 6(2), 559—572.
- Perkal, J. (1967). *Matematyka dla przyrodników i rolników*. Część II. Warszawa: PWN.
- Pluta, W. (1976). Taksonomiczna procedura prowadzenia syntetycznych badań porównawczych za pomocą zmodyfikowanej miary rozwoju gospodarczego. *Przegląd Statystyczny*, 4, 511—517.
- Pociecha, J. (2008). *Rozwój metod taksonomicznych i ich zastosowań w badaniach społeczno-ekonomicznych*. URL: <http://www.stat.gov.pl/cps/rde/xbcr/gus/>.
- Pociecha, J., Podolec, B., Sokołowski, A., Zajac, K. (1988). *Metody taksonomiczne w badaniach społeczno-ekonomicznych*. Warszawa: PWN.
- R Development Core Team (2017). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing. URL: <http://cran.r-project.org>.
- Rusnak, J. (1999). Metoda głównych składowych. W: W. Ostasiewicz (red.), *Statystyczne metody analizy danych*. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu.
- Seidler, J., Badach, A., Molisz, W. (1980). *Metody rozwiązywania zadań optymalizacji*. Warszawa: WNT.
- Strahl, D. (1978). Propozycja konstrukcji miary syntetycznej. *Przegląd Statystyczny*, 2, 205—215.
- Walesiak, M. (1993). Statystyczna analiza wielowymiarowa w badaniach marketingowych. *Prace Naukowe Akademii Ekonomicznej we Wrocławiu*, 654. Seria: *Monografie i Opracowania*, 101.
- Wysocki, F. (2010). *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*. Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu.

- Yazdi, M. M. (2015). *TOPSIS method for multiple-criteria decision making (MCDM) — package topsis*. URL: <https://cran.r-project.org/web/packages/topsis/>.
- Yoon, K. (1980). *System Selection by Multiple Attribute Decision Making*. Rozprawa doktorska. Kansas State University.
- Yoon, K., Hwang, C. L. (1980). *Technique for Order Preference by Similarity to Ideal Solution — A Multiattribute Decision Making*. A paper to be published.
- Yoon, K., Hwang, C. L. (1995). *Multiple Attribute Decision Making: An Introduction*. California: SAGE publications.
- Zeliaś, A. (red.). (2000). *Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym*. Wydawnictwo Akademii Ekonomicznej w Krakowie.

Summary. *The aim of this article is to present the results of multidimensional comparative analysis methods used to assess the state of the environment in Dolnośląskie voivodship in the cross-section of powiats. The research was conducted on the basis of data from the CSO of Poland for 2015 concerning the state and environmental protection in 30 powiats of Dolnośląskie voivodship. The method of linear ordering of objects based on a pattern object (or an anti-pattern object) was used in the research. Many of them described in the subject literature usually lead to differing results (rankings of objects are not the same). It results from i.a. the adopted methods of normalization and weighing of variables and aggregations (creation of synthetic variables). The article is an attempt to compare the results of linear ordering of powiats due the environmental state with the use of method based on a pattern object (or an anti-pattern object). In the rankings correctness analysis, quality indicators were used to evaluate the quality of linear ordering methods.*

Keywords: environmental protection, linear ordering, Hellwig method, TOPSIS method, principal components, R program.