



Adrianna Mastalerz-Kodzis

Uniwersytet Ekonomiczny w Katowicach
Wydział Zarządzania
Katedra Matematyki
adrianna.mastalerz-kodzis@ue.katowice.pl

ZASTOSOWANIE DYNAMICZNEJ WIELOWYMIAROWEJ ANALIZY PORÓWNAWCZEJ W BADANIACH EKONOMICZNYCH

Streszczenie: Celem artykułu jest analiza zmian demograficznych zachodzących w Polsce w latach 2002-2013 w przekroju województw oraz konstrukcja mierników syntetycznych opartych na wielowymiarowej analizie porównawczej z uwzględnieniem dynamiki zmiennych. Za pomocą wybranych metod dynamicznej, wielowymiarowej analizy porównawczej scharakteryzowano województwa w kontekście zmian demograficznych. Na podstawie danych zostały postawione prognozy skonstruowanych mierników oraz został sporządzony ranking województw dla danych historycznych i dla prognoz.

Artykuł składa się z dwóch części. Pierwszą z nich stanowi wprowadzenie metodologiczne dotyczące wielowymiarowej analizy porównawczej, zaś druga część to analiza empiryczna wybranych charakterystyk ekonomicznych, konstrukcja mierników syntetycznych z uwzględnieniem dynamiki zmiennych oraz wnioski.

Słowa kluczowe: wielowymiarowa analiza porównawcza (WAP), demografia, dynamika ekonomiczna.

Wprowadzenie

Starzenie się ludności w krajach Unii Europejskiej w XXI w. jest tematem debat, zarówno w środowisku ekonomistów, jak i polityków. Istnieje wiele różnych ekonomicznych, społecznych i zdrowotnych przyczyn tego zjawiska. Skutki starzenia się społeczeństwa, a przede wszystkim niskiego przyrostu naturalnego, są już widoczne w sferze społecznej. Problem nie dotyczy tylko społeczeństwa polskiego, jednak rozważania ograniczono do danych dotyczących Polski, zaczerpniętych z Banku Danych Lokalnych Głównego Urzędu Statystycznego.

Analiza wybranych wskaźników demograficznych jest zagadnieniem istotnym dla naukowców oraz dla władz państwowych. Niepokojące zmiany w strukturze demograficznej społeczeństwa polskiego, przede wszystkim niski, w wybranych latach nawet ujemny, przyrost naturalny i wzrost frakcji osób w podeszłym wieku, skutkują wieloma zjawiskami społecznymi i ekonomicznymi [Duncan, Scott, 1998; Fihel, Okólski, 2012; Holzer, 2003].

W artykule przedstawiono dynamikę zmian wybranych wielkości ekonomicznych. Pierwszą grupę zmiennych stanowią wybrane charakterystyki demograficzne. Z jednej strony niski przyrost naturalny, z drugiej wydłużający się czas trwania życia powodują istotne zmiany w strukturze wiekowej społeczeństwa polskiego. Nie można jednak rozważać wyłącznie zmiennych demograficznych bez uwzględnienia ich otoczenia społeczno-ekonomicznego w analizie wielowymiarowej. W związku z tym drugą grupę zmiennych stanowią wybrane charakterystyki ogólnoeconomiczne, świadczące o dobrobycie społeczeństwa.

Artykuł jest w znacznej części empiryczny. Jego nadrzędnym celem jest analiza wybranych zmiennych demograficznych, konstrukcja miernika taksonomicznego z uwzględnieniem dynamiki zmiennych oraz zbadanie, czy uwzględnienie dynamiki wpływa na ranking obiektów.

1. Metodologia badań, wybrane elementy WAP

W analizach empirycznych posłużono się wybranymi metodami klasyfikacji danych przestrzenno-czasowych. Opisując n obiektów (województw) za pomocą m zmiennych (charakterystyk ekonomicznych, wskaźników), przekształcono destymulanty i nominanty w stymulanty [Suchecki, red., 2010]. Utworzono macierze danych dla poszczególnych jednostek czasowych (lat).

W klasycznej metodologii wielowymiarowej analizy porównawczej rozważa się dane historyczne. Można wówczas porównać różne obiekty (np. województwa), które są opisywane przez wiele cech, np. przez wskaźniki demograficzne i charakterystyki ekonomiczne. Podczas przekształceń pewna część informacji zostaje jednak utracona, ponieważ stosuje się średnią arytmetyczną charakterystyk. Wszystkie analizowane wielkości zmieniają się pod wpływem czasu, a uśrednienie ich powoduje stratę informacji. W WAP wykorzystuje się wiele wskaźników do oceny poszczególnych obiektów (województw). Na podstawie danych historycznych można prognozować wielkości wskaźników.

Metodologia WAP opiera się na konstrukcji taksonomicznego miernika wzorcowego [Rencher, 2002; Ostasiewicz, red., 1999; Panek, 2009; Suchecki,

red., 2010; Giri, 2004]. Dane o sytuacji województw przedstawia się w postaci macierzy wskaźników. Spośród wielu dostępnych wskaźników należy wybrać kilka najistotniejszych.

Konstrukcja miernika składa się z kilku etapów:

1. Dane historyczne przedstawia się w postaci ciągu macierzy:

$$X(t) = \begin{bmatrix} x_{11}(t) & x_{12}(t) & \dots & x_{1m}(t) \\ x_{21}(t) & x_{22}(t) & \dots & x_{2m}(t) \\ \dots & \dots & \dots & \dots \\ x_{n1}(t) & x_{n2}(t) & \dots & x_{nm}(t) \end{bmatrix} \quad (1)$$

2. Przeprowadza się normalizację zmiennych:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, \quad (2)$$

gdzie \bar{x}_j – średnia arytmetyczna i s_j – odchylenie standardowe cechy j . Macierz znormalizowanych zmiennych jest postaci:

$$Z(t) = \begin{bmatrix} z_{11}(t) & z_{12}(t) & \dots & z_{1m}(t) \\ z_{21}(t) & z_{22}(t) & \dots & z_{2m}(t) \\ \dots & \dots & \dots & \dots \\ z_{n1}(t) & z_{n2}(t) & \dots & z_{nm}(t) \end{bmatrix} \quad (3)$$

3. W macierzy zmiennych znormalizowanych dla każdej zmiennej wybiera się wartość największą; otrzymujemy wówczas wzorec z_{0j} .
4. Oblicza się odległość obiektu od wzorca zgodnie ze wzorem:

$$d_i = \sqrt{\frac{\sum_{j=1}^m (z_{ij} - z_{0j})^2}{m}} \quad (4)$$

Im mniejsza odległość danego obiektu od wzorca, tym mniejsza wartość d_i .

5. Normalizuje się zmienną d_i za pomocą wzoru:

$$z_i = 1 - \frac{d_i}{d_0}, \quad (5)$$

gdzie: z_i – taksonomiczny miernik rozwoju i -tego obiektu,

d_i – odległość i -tego obiektu od wzorca,

d_0 – norma, zmienna z_i przyjmie wartości z przedziału $[0, 1]$ (zastosowano

wzór $d_0 = \bar{d} + 2s_d$, dla \bar{d} – średniej arytmetycznej d_i , s_d – odchylenia standardowego d_i).

Im większa wartość zmiennej syntetycznej z_i , tym obiekt ma wyższą pozycję w rankingu.

W klasycznym ujęciu WAP uśrednia się wartości wskaźników. Badając jednak zmienne o charakterze czasowo-przestrzennych, uwzględniając dynamiczny charakter zmiennych, należałoby prognozować wybrane wielkości. Można prognozować z jednej strony zmienne diagnostyczne [Tarczyński, 2005] i na ich podstawie obliczyć wartości zmiennej syntetycznej; można także prognozować zmienną syntetyczną na podstawie wartości historycznych z_i [Mastalerz-Kodzis, 2014]. W poniżej zapisanej części empirycznej zastosowano klasyczne podejście WAP oraz WAP z uwzględnieniem dynamiki zmiennych.

2. Zastosowanie klasycznej oraz dynamicznej metodologii WAP do analizy danych

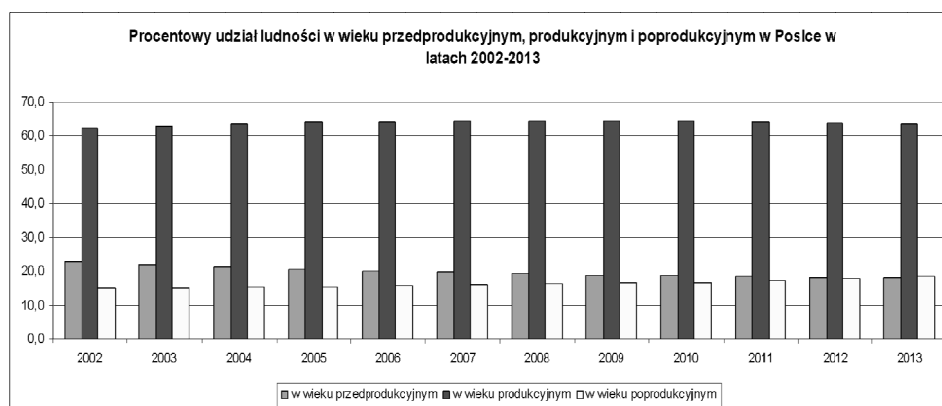
Okresem badawczym są lata 2002-2013. Analizowane są zmienne o charakterze przestrzenno-czasowym, częstotliwości rocznej, z podziałem na województwa. Do analiz posłużyły zmienne zaczerpnięte z BDL GUS:

- odsetek ludności w wieku przedprodukcyjnym,
- odsetek ludności w wieku produkcyjnym,
- odsetek ludności w wieku poprodukcyjnym,
- wskaźnik obciążenia demograficznego,
- przyrost naturalny na 1000 ludności,
- współczynnik dzietności,
- wskaźnik cen towarów i usług,
- wskaźnik aktywności zawodowej,
- stopa bezrobocia,
- przeciętne miesięczne wynagrodzenie brutto.

Powyższe charakterystyki można podzielić na dwie grupy. Pierwszą z nich stanowią wskaźniki demograficzne, drugą wielkości mówiące o aktywności zawodowej ludności i wynagrodzeniach, co świadczy o jakości życia. Wybrano powyższe charakterystyki, ponieważ odsetek młodych ludzi w społeczeństwie

danego województwa implikuje większy przyrost naturalny oraz stymuluje wzrost liczby ludności w regionie, zaś aktywność zawodowa ludności w dłuższej perspektywie zapewnia szybszy i bardziej zrównoważony rozwój regionu, a także korzystnie wpływa na wzrost gospodarczy.

Na rys. 1 przedstawiono procentowy udział ludności w wieku przedprodukcyjnym, produkcyjnym i poprodukcyjnym w Polsce w latach 2002-2013. Można zauważyć, że udział dzieci i młodzieży w wieku przedprodukcyjnym w społeczeństwie w latach 2002-2013 zmalał o 4,5% (22,7% w 2002 r., 18,2% w 2013 r.), zaś udział osób w wieku poprodukcyjnym wzrósł o 3,3% (15,1% w 2002 r., 18,4% w 2013 r.). Jest to konsekwencją niskiego, nawet ujemnego, przyrostu naturalnego oraz wydłużającego się czasu trwania życia.



Rys. 1. Struktura wieku ludności w Polsce w latach 2002-2013

Na podstawie ciągu macierzy danych sporządzono ranking województw z zastosowaniem klasycznej metodologii WAP dla danych historycznych, uśredniając wielkości, nie uwzględniając dynamiki zmian. Ranking województw przedstawiono w kolumnie 3 tab. 1. W dalszej kolejności, uwzględniając dane historyczne za lata 2002-2013 dla każdej zmiennej i dla każdego województwa, wyznaczono metodą trendu liniowego prognozy zmiennych na lata 2014, 2015. Dopasowanie linii trendu w każdym przypadku było na poziomie 0,72-0,93. Następnie dla prognoz wyznaczono wartości miary syntetycznej z_i i ustalono pozycje województw. Zmieniły się wartości miernika syntetycznego. Wyniki zawarto w kolumnie 2 i 4 tab. 1. W obliczeniach stosowano takie same wagi dla poszczególnych zmiennych diagnostycznych.

Tab. 1. Ranking województw dla danych historycznych na podstawie wartości miary syntetycznej z_i

Województwo	Z_i (dane historyczne 2002-2013)	Pozycja województwa w rankingu, dane 2002-2013
DOLNOŚLĄSKIE	0,1247	13
KUJAWSKO-POMORSKIE	0,2994	5
LUBELSKIE	0,2865	7
LUBUSKIE	0,2832	6
ŁÓDZKIE	0,0592	16
MAŁOPOLSKIE	0,3702	4
MAZOWIECKIE	0,4417	2
OPOLSKIE	0,0682	15
PODKARPACKIE	0,2275	10
PODLASKIE	0,2569	9
POMORSKIE	0,5057	1
ŚLĄSKIE	0,1992	12
ŚWIĘTOKRZYSKIE	0,0969	14
WARMIŃSKO-MAZURSKIE	0,2742	8
WIELKOPOLSKIE	0,4407	3
ZACHODNIOPOMORSKIE	0,2194	11

Dla danych historycznych można stwierdzić, że biorąc pod uwagę wybrane zmienne demograficzne oraz wybrane charakterystyki aktywności zawodowej ludności i ich wynagrodzeń najlepsza sytuacja jest w województwach kolejno: pomorskim, mazowieckim i wielkopolskim, zaś najbardziej niepokojąca w województwach: łódzkim, opolskim, świętokrzyskim i dolnośląskim. Najwyżej w rankingu są województwa o wysokim wskaźniku aktywności zawodowej, niskiej stopie bezrobocia i największym przeciętnym wynagrodzeniu brutto. W tych województwach jest także największy odsetek ludności w wieku produkcyjnym. Największe różnice w zmiennych: przeciętne miesięczne wynagrodzenia brutto i stopa bezrobocia notuje się w przypadku województw będących na najwyższych i najniższych pozycjach w rankingu.

Następnie sporządzono prognozy zmiennych diagnostycznych i na podstawie prognoz na lata 2014 i 2015 wyznaczono wartości zmiennej syntetycznej. Wyniki oraz uszeregowanie województw przedstawia tab. 2.

Tab. 2. Ranking województw dla prognoz zmiennych diagnostycznych na lata 2014, 2015

Województwo	Z_i na podstawie prognoz zmiennych diagnostycznych 2014	Pozycja w rankingu na podstawie prognoz zmiennych diagnostycznych 2014	Z_i na podstawie prognoz zmiennych diagnostycznych 2015	Pozycja w rankingu na podstawie prognoz zmiennych diagnostycznych 2015
DOLNOŚLĄSKIE	0,2364	10	0,2473	10
KUJAWSKO- -POMORSKIE	0,2819	9	0,2752	9
LUBELSKIE	0,2935	7	0,2882	7
LUBUSKIE	0,3695	5	0,3743	5
ŁÓDZKIE	0,1504	14	0,1592	14
MAŁOPOLSKIE	0,4162	4	0,4118	4
MAZOWIECKIE	0,5916	1	0,5860	1
OPOLSKIE	0,0871	15	0,0962	15
PODKARPACKIE	0,2166	12	0,2038	13
PODLASKIE	0,2828	8	0,2811	8
POMORSKIE	0,5655	2	0,5660	2
ŚLĄSKIE	0,3150	6	0,3302	6
ŚWIĘTOKRZYSKIE	0,0591	16	0,0499	16
WARMIŃSKO- -MAZURSKIE	0,2326	11	0,2184	11
WIELKOPOLSKIE	0,4935	3	0,4870	3
ZACHODNIOPOMORSKIE	0,2139	13	0,2117	12

Po uwzględnieniu prognoz na lata 2014, 2015 można zauważyć niewielkie różnice pomiędzy pozycjami województw. Zgodnie z wartościami miernika syntetycznego w latach 2014, 2015 nadal najlepiej pod względem demograficzno-ekonomicznym mieszkać się będzie w województwach: mazowieckim, pomorskim i wielkopolskim. Najgorsza sytuacja będzie dotyczyć województw: świętokrzyskiego, opolskiego, łódzkiego, zachodniopomorskiego i podkarpackiego.

Ponadto pozycje w rankingu kilku województw znacznie zmieniły się po uwzględnieniu prognoz na lata 2014, 2015. Na niekorzyść zmienił się ranking dla województw: kujawsko-pomorskiego, podkarpackiego, świętokrzyskiego i warmińsko-mazurskiego. W tych województwach poziom życia ludności obniżył się oraz można zauważyć niekorzystny trend zmian demograficzno-ekonomicznych. Korzystnie w rankingu zmieniły się natomiast pozycje województw śląskiego i dolnośląskiego. Świadczy to o korzystnym trendzie zmian w sytuacji życia ludności tych województw.

Współczynnik korelacji rang Spearmana dla rankingów dla danych historycznych i dla prognozowanych na 2014 r. jest równy 0,867647. Kolejność województw w znacznym stopniu nie uległa zatem zmianie. Współczynnik korelacji

rang dla rankingów w latach 2014 i 2015 wynosi 0,997059. Istnieje bardzo silna zależność pomiędzy wartościami z_i w omawianych latach. Jest to wynikiem zastosowania funkcji regresji do modelowania zmiennych diagnostycznych.

Zastosowano także drugi wariant, a mianowicie na podstawie ciągu historycznych wartości z_i w poszczególnych latach obliczono za pomocą regresji liniowej wartości miary syntetycznej w latach 2014, 2015. Wyniki przedstawiono w tab. 3.

Tab. 3. Wartości miary syntetycznej dla prognoz na lata 2014, 2015 wraz z rankingiem województw

Województwo	Prognoza z_i na 2014 r.	Pozycja w rankingu według prognozy z_i na 2014 r.	Prognoza z_i na 2015 r.	Pozycja w rankingu według prognozy z_i na 2015 r.
DOLNOŚLĄSKIE	0,2245	11	0,2367	10
KUJAWSKO-POMORSKIE	0,2662	9	0,2585	9
LUBELSKIE	0,3107	7	0,3132	7
LUBUSKIE	0,3483	5	0,3585	5
ŁÓDZKIE	0,1422	14	0,1528	14
MAŁOPOLSKIE	0,4441	4	0,4558	4
MAZOWIECKIE	0,6476	1	0,6770	1
OPOLSKIE	0,0771	15	0,0770	15
PODKARPACKIE	0,2062	12	0,2002	13
PODLASKIE	0,2980	8	0,3036	8
POMORSKIE	0,5659	2	0,5754	2
ŚLĄSKIE	0,3166	6	0,3338	6
ŚWIĘTOKRZYSKIE	0,0674	16	0,0579	16
WARMIŃSKO-MAZURSKIE	0,2246	10	0,2180	11
WIELKOPOLSKIE	0,4715	3	0,4747	3
ZACHODNIOPOMORSKIE	0,2041	13	0,2012	12

Z powyższych analiz można wysunąć następujące wnioski. Wartości zmiennej syntetycznej różnią się w zależności od zastosowanej metodologii. Można jednak stwierdzić, że wiodące województwa według tab. 1 są także na najwyższych pozycjach w rankingu w tab. 2 i 3 zarówno w 2014 r., jak i 2015 r. Uwzględnienie zatem dynamiki, czy to zmiennych diagnostycznych, czy też wartości miernika syntetycznego, daje porównywalne rezultaty. Ponadto współczynniki korelacji Spearmana pomiędzy rankingami są bliskie jedności. W 2014 r. współczynnik ten wynosił 0,997059, zaś w 2015 r. jest równy 1. Istnieje zatem bardzo silna korelacja pomiędzy rankingami. Można stwierdzić, że pozycja w rankingu nie zależy od zastosowanej metody prognozy wartości miary syntetycznej z_i .

Podsumowanie

Konkludując, można zapisać, że wartości zmiennej syntetycznej różnią się w zależności od zastosowanej metodologii. Uwzględnienie dynamiki, zarówno zmiennych diagnostycznych, jak i wartości miernika syntetycznego, daje zbliżone wyniki. Pozycja obiektu w rankingu nie zależy od zastosowanej metody prognozowania przyszłej wartości miary syntetycznej.

Złożoność otaczającego świata wymaga stosowania do jego opisu metod wielowymiarowych, uwzględniających dynamikę zmiennych. W analizach ekonomicznych często wykorzystuje się metodykę WAP, jednak w niewielkim stopniu uwzględniając dynamiczny charakter zmian. Analiza wybranych wielkości demograficznych z uwzględnieniem zmiennych ogólnoeconomicznych pozwala na skonstruowanie mierników syntetycznych umożliwiających rankingowanie regionów ze względu na badane wielkości i wskazanie tych regionów, w których sytuacja jest najlepsza oraz pokazanie tych województw, gdzie jest najgorzej.

Z uwagi na zmieniające się w czasie wielkości zmiennych diagnostycznych, konieczne jest ciągle monitorowanie sytuacji, w szczególności za pomocą metod ilościowych, aby odpowiednio wcześniej reagować na nadchodzące zmiany.

Literatura

- Duncan C.J., Scott S. (1998), *Human Demography and Disease*, University Press, Cambridge, s. 103-110.
- Fihel A., Okólski M. (2012), *Demografia. Współczesne zjawiska i teorie*, Scholar, Warszawa.
- Giri C.N. (2004), *Multivariate Statistical Analysis*, USA Marcel Dekker, s. 435-526.
- Holzer J.Z. (2003), *Demografia*, PWE, Warszawa.
- Mastalerz-Kodzis A. (2014), *Konstrukcja optymalnych portfeli z zastosowaniem metod analizy fundamentalnej – ujęcie dynamiczne*, „Studia Ekonomiczne”, nr 163, Uniwersytet Ekonomiczny, Katowice.
- Ostasiewicz W., red. (1999), *Statystyczne metody analizy danych*, Wydawnictwo AE, Wrocław.
- Panek T. (2009), *Statystyczne metody wielowymiarowej analizy porównawczej*, Oficyna Wydawnicza SGH, Warszawa.
- Rencher A.C. (2002), *Methods of Multivariate Analysis*, John Wiley & Sons, USA.
- Suhecki B., red (2010), *Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych*, C.H. Beck, Warszawa.
- Tarczyński W. (2005), *Dynamiczne ujęcie taksonomicznej miary atrakcyjności inwestowania na przykładzie wybranych spółek notowanych na Gieldzie Papierów Wartościowych w Warszawie [w:] Metody Ilościowe w Ekonomii, cz. I, „Zeszyty Naukowe Uniwersytetu Szczecińskiego”, nr 394.*

**ANALYSIS OF CHOSEN ECONOMICAL CHARACTERISTICS BY
DYNAMICAL MULTIVARIATE COMPARATIVE ANALYSIS**

Summary: The aim of the paper is to analyze the dynamic demographic and economical changes that are observed in Poland from 2002 to 2013. General tendencies in Polish population age structure and their impact on selected labour indicators were shown. Besides, some dependencies between demographic variables and chosen economic indicators were presented. By means of selected multivariate comparative analysis tools (the synthetic measure of development) the regions of Poland were characterized and compared.

Keywords: multivariate comparative analysis, demography, dynamical economic.