


WYKORZYSTANIE WSPÓLCZYNNIKA MORANA DO KLASYFIKACJI OBIEKTÓW PRZESTRZENNYCH NA PRZYKŁADZIE CEN GRUNTÓW ROLNICZYCH

Robert Pietrzykowski  <https://orcid.org/0000-0002-6122-3043>
Instytut Ekonomii i Finansów
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: robert_pietrzykowski@sggw.edu.pl

Streszczenie: W artykule przedstawiono możliwości wykorzystania globalnej statystyki Morana do klasyfikacji obiektów przestrzennych. Przestrzenna autokorelacja jest wykorzystywana przede wszystkim do określenia podobieństwa obiektów, które znajdują się w różnych odległościach od siebie. Sąsiedztwo badanych obiektów, określamy przez jego rząd czyli obiekty, które sąsiadują bezpośrednio ze sobą traktujemy jako sąsiadów pierwszego rzędu. Natomiast dalsi sąsiedzi określani są jako sąsiedztwo drugiego, trzeciego i dalszych rzędów. Współczynniki autokorelacji globalne i lokalne pozwalają określić wielkość zależności przestrzennej, ale służą przede wszystkim do wnioskowania o strukturze przestrzennej. Wykorzystując globalne miary statystyki przestrzennej możemy określić istnienie reżimów przestrzennych, a w konsekwencji klasyfikować badane obiekty w klastry.

Słowa kluczowe: autokorelacja przestrzenna, ceny gruntów rolniczych

JEL classification: C14, R3

WSTĘP

Współczynniki autokorelacji przestrzennej służą do oceny korelacji zmiennych w odniesieniu do lokalizacji przestrzennych. Przestrzenna autokorelacja w rozumieniu geograficznym określa podobieństwo obiektów, które znajdują się w różnych odległościach od siebie. Współczynniki autokorelacji pozwalają określić wielkość zależności przestrzennej, ale służą przede wszystkim do wnioskowania o strukturze przestrzennej. Wykorzystując miary statystyki przestrzennej możemy określić istnienie reżimów przestrzennych a w konsekwencji klasyfikować badane

<https://doi.org/10.22630/MIBE.2020.21.3.16>

obiekty w klastry. Do globalnych statystyk autokorelacji zaliczamy: współczynnik Morana, C-Geara i Joint-count. Natomiast do lokalnych: współczynnik lokalny Morana, a także wchodzący w skład tzw. LISA (ang. Local Indicator of Spatial Association, Anselin 1995) CL-Geara oraz statystyka G-Getisa [Ord i Getis 1995]. Poszerzeniem miary lokalnej G jest zaproponowana przez Getisa i Orda [2012] miara LOSH (ang. Local spatial heteroscedansity), która pozwala na bardziej szczegółowe analizy przestrzenne.

Celem pracy było pokazanie możliwości wykorzystania globalnej statystyki Morana do klasyfikacji obiektów przestrzennych na przykładzie cen gruntów rolniczych.

BADANIE ZALEŻNOŚCI

Ze względu na rodzaj obserwowanych zmiennych oraz charakteryzujących je rozkładów do badania zależności wykorzystuje się zwykle współczynniki korelacji Pearsona, Spearmana, Kendalla i V-Cramera. Klasyczny współczynnik korelacji Pearsona jest miarą liniowej zależności i dotyczy analizy cech ciągłych, które charakteryzują się dwuwymiarowym rozkładem normalnym. Do badania wielowymiarowej normalności można wykorzystać test zaproponowany przez Mardia w 1970 r., lub jego modyfikację zaproponowaną przez Hanusz i Tarasińską w 2014. Przy niespełnieniu tych założeń należy wykorzystać rangowe współczynniki korelacji lub testy nieparametryczne.

W analizie przestrzennej do badania korelacji wykorzystuje się dwa typy miar, a mianowicie miary globalne i lokalne. Miary globalne są wskaźnikiem autokorelacji przestrzennej, który dotyczy ogólnego podobieństwa regionów. Natomiast statystyki lokalne pozwalają na określenie podobieństwa badanego regionu w stosunku do jego sąsiadów. Punktem wyjścia konstrukcji globalnych współczynników korelacji oprócz ogólnej koncepcji liniowego współczynnika korelacji Pearsona jest statystyka gamma, która łączy ze sobą informacje dotyczące macierzy odległości między badanymi obiektami i macierzy korelacji między wartościami badanej zmiennej w poszczególnych lokalizacjach. Statystyka gamma nazywana jest również indeksem Mantela [1967], który jako pierwszy opracował i zastosował test korelacji między dwiema macierzami w zagadnieniach epidemiologicznych. W literaturze zagadnienie to znane jest również jako „paradygmat analiz kwadratowych” QAP (ang. Quadratic Analysis Paradigm) [Hubert i in. 1985, Anselin 1992]. W ogólnej postaci statystykę gamma zapiszemy jako:

$$\Gamma = \mathbf{WR} = \sum_{i=1}^N \sum_{j=1}^N w_{ij} r_{ij}, \quad (1)$$

gdzie: \mathbf{W} jest macierzą odległości, a \mathbf{R} jest macierzą korelacji obserwowanych zmiennych. Statystyka gamma jest funkcją iloczynów krzyżowych miar odległości w badanej przestrzeni (w_{ij}) i miar współzależności (r_{ij}) wartości zmiennej Y. Chcąc przedstawić globalną statystykę Morana w postaci miary typu QAP najlepiej dokonać standaryzacji wartości zmiennych zgodnie z formułą:

$$z_i = \frac{y_i - \bar{y}}{s_y}, \quad (2)$$

co w konsekwencji prowadzi do uzyskania statystyki QAP Morana, którą zapisujemy jako:

$$\Gamma_M = \sum_{i=1}^N \sum_{j=1}^N w_{ij} (z_i \cdot z_j), \quad (3)$$

Ostateczna postać statystyki Morana różni się od statystyki QAP Morana (formuła 3) ponieważ jej postać uzyskuje się przez podzielenie konkretnej statystyki przez wariancję z próby w celu uzyskania odpowiedniego przedziału liczbowego dla miernika korelacji (-1; 1):

$$M_{QAP} = \frac{\Gamma_M}{n \cdot \text{var}(Y)} = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (z_i \cdot z_j)}{n \cdot \text{var}(Y)} = \frac{\mathbf{z}' \mathbf{W} \mathbf{z}}{\mathbf{z}' \mathbf{z}} \quad (4)$$

gdzie: \mathbf{z} jest wektorem kolumnowym o elementach $z_i = y_i - \bar{y}$, kiedy rozpatrujemy statystykę Morana dla zmiennych Y , które są zbiorem dotyczącym obserwacji w wielu lokalizacjach y_i dla ($i = 1, \dots, N$).

Mając niestandardyzowaną macierz odległości \mathbf{W} oraz nietransformowane wartości obserwacji, statystyka Morana obliczana z zastosowaniem miar podobieństwa w postaci iloczynów krzyżowych odchyłeń od średniej ma postać:

$$M_{gl} = \frac{1}{\sum_i^n \sum_j^n w_{ij}} \cdot \frac{\sum_i^n \sum_j^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\frac{1}{N} \sum_i^n (y_i - \bar{y})^2} = \frac{N}{S_0} \cdot \frac{\mathbf{z}' \mathbf{W} \mathbf{z}}{\mathbf{z}' \mathbf{z}}, \quad (5)$$

gdzie: $S_0 = \sum_i^n \sum_j^n w_{ij}$ jest sumą wszystkich elementów macierzy \mathbf{W} .

WYKORZYSTANE DANE

Dane wykorzystane w pracy pochodziły z Systemu Analiz i Monitorowania Rynku Obrotu Nieruchomościami (AMRON). Zgromadzone dane dotyczą ich lokalizacji i cen transakcyjnych z aktów notarialnych, wartości wycen, ofert lub innych w pełni weryfikowalnych źródeł. Baza spełnia najwyższe standardy jakości zdefiniowane przez Komisję Nadzoru Finansowego. W analizach wykorzystano ceny gruntów rolniczych na poziomie powiatów i gmin z lat 2004 – 2012. Dokładny opis danych przedstawił Pietrzykowski w monografii „Przestrzenne różnicowanie cen ziemi rolniczych w Polsce” [2019].

METODA BADAWCZA

W pracy wykorzystano globalny współczynnik autokorelacji Morana [1948] do klasyfikacji badanych obiektów. Globalny współczynnik autokorelacji Morana określa zależności uwzględniające wzajemne położenie obiektów. Dla danych przestrzennych mówimy zatem o opóźnieniu przestrzennym, które jest spowodowane przez kryterium sąsiedztwa, a tym samym powoduje autokorelację przestrzenną. Konsekwencją istnienia zależności przestrzennej jest grupowanie się

podobnych wartości w klastry, co wykorzystano w badaniach. Związki przestrzenne mogą wynikać z istnienia korelacji w obrębie całego badanego obszaru i wtedy mówimy o globalnej korelacji. Chcąc stwierdzić, czy badane obiekty są rozmieszczone w całkowicie losowy sposób pomiędzy poszczególnymi lokalizacjami należy zweryfikować hipotezę o braku autokorelacji przestrzennej o postaci:

$$H_0: \rho = 0, H_1: \rho \neq 0, \quad (6)$$

dla dostatecznie dużych N unormowana statystyka Morana $Z(M_{gl})$ o postaci:

$$Z(M_{gl}) = \frac{M_{gl} - E(M_{gl})}{\sqrt{\text{var}(M_{gl})}} \sim N(0; 1), \quad (7)$$

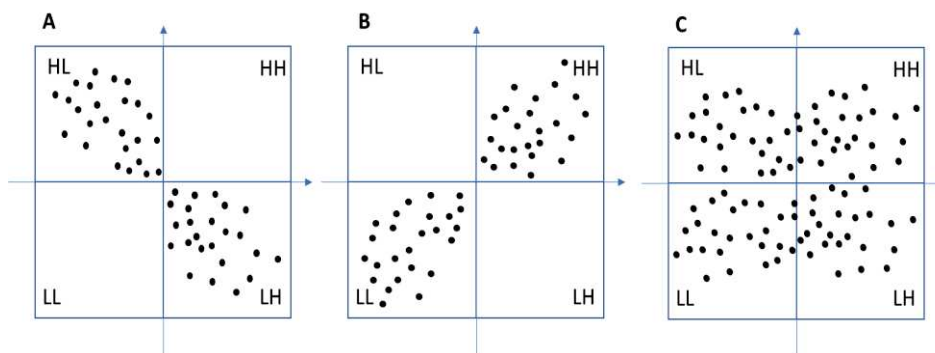
ma rozkład normalny o wartości oczekiwanej 0 i wariancji 1 [Cliff i Ord 1973]. Odrzucenie hipotezy zerowej na rzecz hipotezy alternatywnej (H_1) oznacza istnienie autokorelacji przestrzennej i implikuje wniosek, że wartości obserwowanej zmiennej nie są rozmieszczone w sposób losowy pomiędzy poszczególnymi lokalizacjami. Do weryfikacji hipotezy zerowej wykorzystuje się globalny współczynnik korelacji Morana zapisany formułą 5. W pracy wykorzystano macierz wag pierwszego rzędu określającą jedynie najbliższych sąsiadów:

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1j} \\ w_{21} & w_{22} & \dots & w_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ w_{i1} & w_{i2} & \dots & w_{ij} \end{bmatrix}, \quad (8)$$

gdzie: poszczególne elementy macierzy uzyskano wstawiając $w_{ij} = 1$ w przypadku kiedy obiekt i -ty był sąsiadem obiektu j -tego. Natomiast $w_{ij} = 0$, kiedy obiekt i -ty nie był sąsiadem obiektu j -tego oraz na diagonalu macierzy \mathbf{W} , ponieważ dany obiekt nie może być swoim własnym sąsiadem. Tak zapisana macierz pierwszego rzędu nazywana jest również macierzą binarną. Współczynnik korelacji Morana i diagram Morana posłużył do określenia struktury związków przestrzennych. Budując diagram Morana dane oryginalne poddaje się standaryzacji. Na osi x -ów występuje zatem standaryzowana wartość cechy, a na osi y -ów opóźnienie przestrzenne. Diagram Morana dzielimy na cztery części względem wartości zerowych. Rozłożenie punktów w przestrzeni wykresu świadczy o obserwowanej autokorelacji. W prowadzonych badaniach zostało to wykorzystane do określenia i podziału obiektów (gmin) ze względu na badaną cechę na klastry. W celu poprawienia czytelności wyników, tak uzyskane grupy obiektów w diagramach Morana przeniesiono na mapę Polski (patrz rysunek 2). Rozłożenie punktów na wykresie Morana (rysunek 1), świadczy o obserwowanej autokorelacji i tak: punkty w kwadracie HL i LH wskazują na ujemną autokorelację, a w kwadratach LL i HH na dodatnią autokorelację. Równomierne rozłożenie punktów w czterech kwadratach świadczy o braku autokorelacji przestrzennej. Jak wspomniano wcześniej wykres Morana wykorzystano do określenia klastrów, które uzyskano odczytując rozmieszczenie punktów w poszczególnych kwadratach wykresu.

Obiekty, które znalazły się w kwadracie LL i HH ze względu na badaną cechę w tym regionie przyjmują odpowiednio niskie lub wysokie wartości. Lewa kolumna obejmująca kwadraty HL i LL, zawiera wartości niskie w regionach sąsiadujących, a kolumna prawa zawierająca kwadraty HH i LH, dotyczy wysokich wartości cechy w regionach sąsiadujących. Analogicznie górny wiersz (kwadraty HL, HH) zawiera wartości wysokie w regionie i-tym, a wiersz dolny (kwadraty LL, LH) obejmuje wartości niskie w regionie i-tym [Anselin 2001, Kopczewska 2020].

Rysunek 1. Przykładowe diagramy Morana obrazujące występowanie ujemnej (A), dodatniej (B) i brak (C) autokorelacji przestrzennej



Źródło: opracowanie własne

W celu potwierdzenia grupowania się obiektów w układy klastrów wykorzystano test Kruskala—Wallisa [Kruskal 1952, Kruskal i Wallis 1952], który jest nieparametryczną wersją jednoczynnikowej analizy wariancji. Podobnie jak w jednoczynnikowej analizie wariancji (ANOVA), weryfikuje się hipotezę o równości średnich z tą różnicą, że zamiast średnich bierzemy pod uwagę mediany:

$$H_0: \theta_1 = \theta_2 = \dots = \theta_k,$$

$$H_1: \text{nie wszystkie } \theta_j \text{ są sobie równe } (j = 1, 2 \dots k) \quad (9)$$

gdzie: $\theta_1, \theta_2, \dots, \theta_k$ to mediany badanej zmiennej w populacji. Statystykę testową zapisujemy następującą formułą:

$$KW_{emp} = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{(\sum_{i=1}^{n_j} G_{ij})^2}{n_i} - 3(N+1) \quad (10)$$

gdzie: G_{ij} określa rangi przypisane do wartości zmiennej dla $i = 1, 2, \dots, n_j, j = 1, 2, \dots, k$. Jedynym ograniczeniem do stosowania tego testu jest wielkość próby (dla $n_j > 5$). Zaletą testu Kruskala – Wallisa jest to, że nie wymaga on spełnienia założeń dotyczących normalności rozkładów, a także stałości wariancji. W tym przypadku zakładamy, że badane populacje różnią się jedynie parametrem położenia, a zatem mają taki sam kształt i parametr skali. W przypadku odrzucenia hipotezy zerowej (formuła 9) należy zastosować procedurę porównań szczegółowych dla median

stosując testy porównań szczegółowych np. test Dunna [1964], Conovera-Imana [1999] lub Wilcoxona [1945]:

$$H_0: \theta_j = \theta_{j+1}, H_1: \theta_j \neq \theta_{j+1}, \quad (11)$$

Do weryfikacji serii hipotez (formuła 11) wykorzystano test Wilcoxona [Conover 1999] o postaci:

$$T = \min\{\sum G(-), \sum G(+)\}, \quad (12)$$

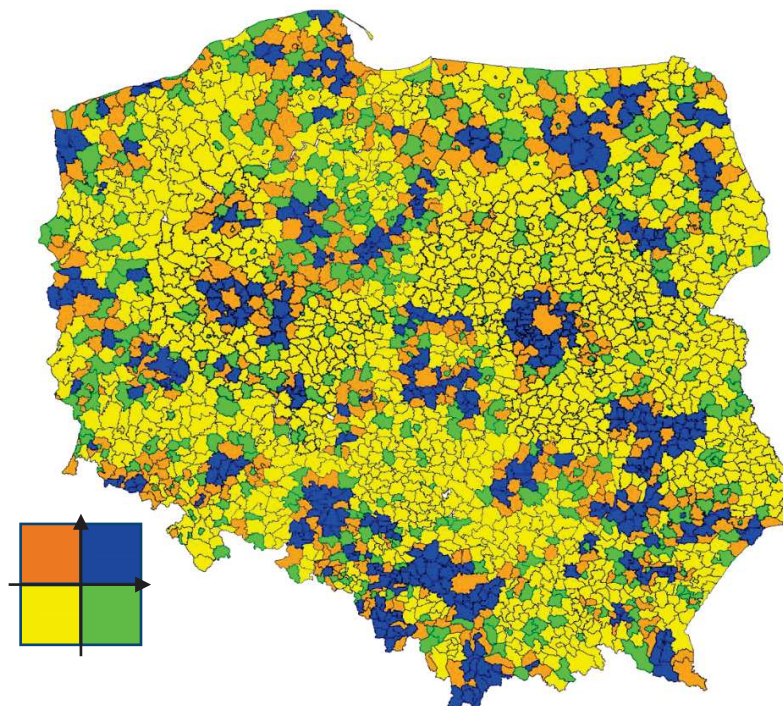
gdzie: $\sum G(-)$, $\sum G(+)$ oznaczają sumy rang ujemnych i dodatnich dla badanych prób. Przy prawdziwości hipotezy zerowej (formuła 11) statystyka T ma rozkład Wilcoxona, który przy dużych próbach można aproksymować testem t Studenta [Wilcoxon 1945].

WYNIKI BADAŃ

W pracy prowadzono badania obiektów zgodnie z klasyfikacją NUTS, która została wprowadzona w Polsce od 26 listopada 2005 roku. Poziom NUTS1 dzieli Polskę na makroregiony grupujące województwa. NUTS2 dzieli Polskę na województwa, a NUTS3 dzieli Polskę na podregiony grupujące powiaty. Rozporządzenie NUTS pozwala na wydzielenie lokalnych jednostek administracyjnych (LAU¹), które są różne dla poszczególnych państw Unii Europejskiej. W Polsce gminy są zatem lokalnym poziomem jednostek administracyjnych LAU, na bazie których przeprowadzono analizy z wykorzystaniem statystyki Morana. Na rysunku 2 przedstawiono klasyfikację cen ziemi rolniczej w gminach w Polsce w roku 2004 na podstawie statystyki i diagramu Morana. Poszczególne kolory oznaczają odpowiednio ćwiartki diagramu Morana do których zostały przyporządkowane poszczególne obiekty (gminy). Kolor pomarańczowy oznacza gminy o niskich cenach ziemi otoczonych przez gminy w których zaobserwowano wysokie ceny gruntów (LH). Kolor niebieski oznacza te gminy, w których wartości cen gruntów są zarówno wysokie w regionie, jak i w jego otoczeniu (HH). Kolor żółty to gminy o niskich średnich wartościach w regionie jak i w jego otoczeniu. Kolor zielony to gminy o średnich niskich wartościach w regionie i wysokich w otoczeniu. Zauważmy, że duże miasta wojewódzkie są otoczone przez pierścienie (gminy) w których zaobserwowano wysokie ceny gruntów. W miarę oddalania się od centrum miasta średnie ceny ziemi w gminach są coraz niższe. Podobny podział na klastry uzyskano dla pozostałych lat tzn. od roku 2005 do 2012. W roku 2012 nie obserwowano tak silnego zarysowania pierścieni wokół miast, ale mogły mieć na to wpływ braki w danych.

¹ <https://ec.europa.eu/eurostat/web/nuts/local-administrative-units>

Rysunek 2. Klasyfikacja cen ziemi rolniczej w gminach w Polsce w roku 2004 na podstawie statystyki i diagramu Morana

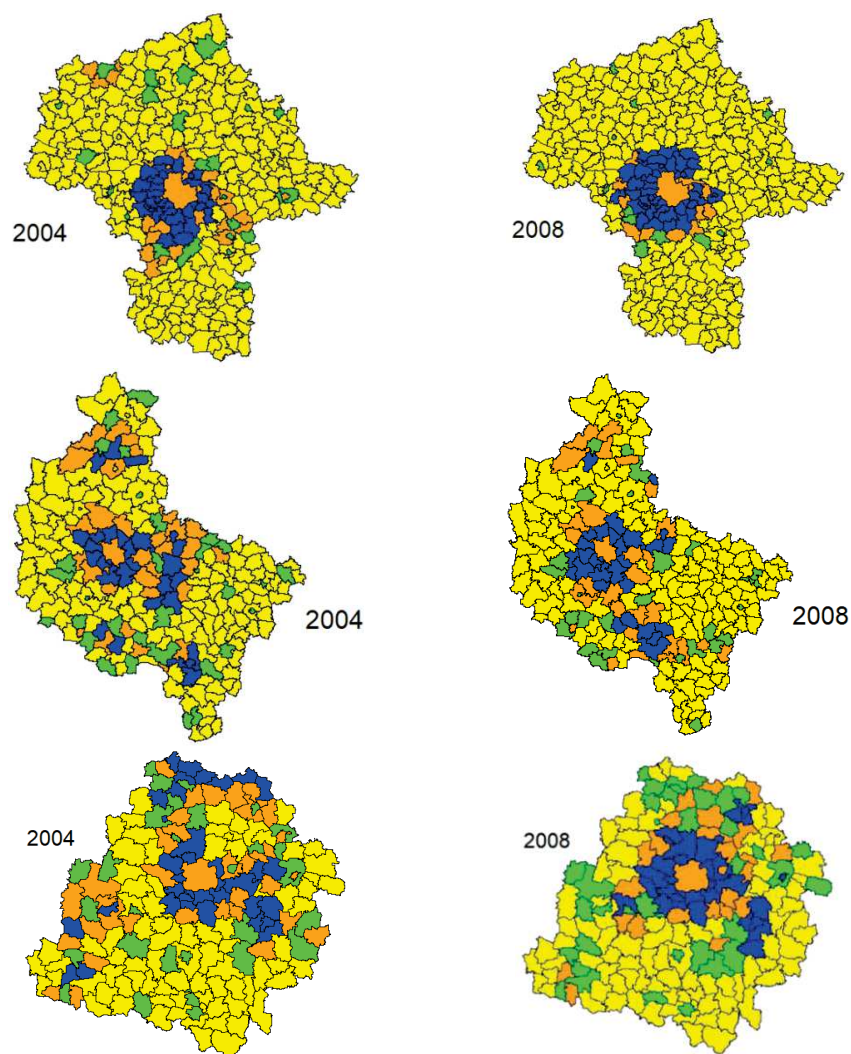


Źródło: opracowanie własne

Szczegółową analizę przeprowadzono dla wszystkich województw z podziałem na gminy. We wszystkich województwach uzyskano podobne wyniki grupowania gmin tzn. duże miasto wojewódzkie silnie oddziaływało na grupowanie badanych jednostek. Znajdowało to swoje potwierdzenie w poszczególnych latach badanego okresu tzn. 2004 – 2012. W niektórych województwach obserwowano podobne podziały dla mniejszych miast. Jednak najsilniejszy podział i oddziaływania miało duże miasto wokół, którego grupowały się gminy. W dalszej części rozdziału przedstawiono, jak wyglądało szczegółowe badanie dla poszczególnych gmin na przykładzie województwa mazowieckiego, łódzkiego i wielkopolskiego. Na rysunku 3 zamieszczono klasyfikację cen ziemi rolniczej w gminach w Polsce w roku 2004 i 2008 na podstawie statystyki i diagramu Morana w trzech województwach. Na rysunku 3 patrząc od góry mamy województwo mazowieckie, następnie w środku wielkopolskie i na dole łódzkie. Analizy dla innych polskich województw wyglądały podobnie, jednak ze względu na ograniczoną objętość pracy przedstawiono jedynie te trzy województwa. Przedstawione lata 2004 i 2008 nie są przypadkowe. Rok 2004 można traktować jako graniczny ze względu na przystąpienie Polski do Unii Europejskiej. Natomiast

rok 2008 jest czasem w którym wystąpił krach gospodarczy na świecie. Zauważmy, że nie miał on wpływu na zmiany cen gruntów rolniczych. Można stwierdzić, że podziały w roku 2008 są silniejsze niż w roku 2004.

Rysunek 3. Klasyfikacja cen ziemi rolniczej w gminach w Polsce w roku 2004 i 2008 na podstawie statystyki i diagramu Morana w województwach: mazowieckim, wielkopolskim i łódzkim



Źródło: opracowanie własne

W dalszej części pracy szukano potwierdzenia uzyskanych klasyfikacji cen ziemi rolniczej w gminach w Polsce na podstawie statystyki i diagramu Morana.

W tym celu wykorzystano test Kruskala–Wallisa i test Wilcozona. Jak wyjaśniono wcześniej, wykorzystanie klasycznej analizy wariancji nie było możliwe ze względu na niespełnienie założeń, zarówno dotyczących równości wariancji, jak i rozkładu normalnego. Dla poszczególnych podziałów weryfikowano następujące hipotezę:

$$H_0: \theta_1 = \theta_2 = \theta_3 = \theta_4, \quad (13)$$

gdzie: $\theta_1, \theta_2, \theta_3, \theta_4$ oznaczają mediany dotyczące kolejnych klastrów (gmin, dla poszczególnych województw). Hipotezę zapisaną formułą 13 odrzucono dla każdego z badanych województw, co pozwoliło potwierdzić, że uzyskany podział nie jest losowy. W celu wydzielenia grup różniących się między sobą zastosowano test Wilcozona z uwzględnieniem poprawki na wielokrotne testowanie. Przykładowe wyniki porównań szczegółowych dla poszczególnych grup od 1 do 4, dla województwa mazowieckiego zapisano w tabeli 1. Wartości p-value zamieszczone w tabeli, dotyczą porównań szczegółowych dla badanych median, reprezentujących odpowiednie klastry w województwie mazowieckim.

Tabela 1. Wyniki podziału na grupy homogeniczne dla cen gruntów rolniczych w województwie mazowieckim w roku 2004, potwierdzające brak losowości w podziałach uzyskanych na podstawie współczynnika korelacji i diagramu Morana

Hipotezy	p-value	Mediany	Oznaczenie na wykresie	Grupy homogeniczne
$H_0: \theta_1 = \theta_2$	< 2.2e-16	θ_1	Blue HH	A
$H_0: \theta_1 = \theta_3$	< 2.2e-16	θ_2	Green HL	B
$H_0: \theta_1 = \theta_4$	< 2.2e-16	θ_3	Yellow LL	C
$H_0: \theta_2 = \theta_3$	< 2.2e-16	θ_4	Orange LH	C
$H_0: \theta_2 = \theta_4$	< 2.2e-16			
$H_0: \theta_3 = \theta_4$	0,2657			

Źródło: opracowanie własne.

Na podstawie przeprowadzonych badań (tabela 1) można stwierdzić występowanie różnic pomiędzy prawie wszystkimi grupami (wartości p-value są mniejsze od zadanego poziomu istotności 0,05). Nie udało się rozróżnić jedynie grup oznaczonych jako 3 i 4, czyli skupień (reżimów przestrzennych) znajdujących się w trzeciej i czwartej ćwiartce ($0,05 < 0,2657$). Można zatem wyciągnąć wniosek, że najwyższe ceny ziemi lokalizują się wokół Warszawy, po czym następuje ich zmniejszanie. Struktura ta utrzymywała się w badanym okresie tzn od 2004 do 2012 roku. Analogiczne badania przeprowadzono dla pozostałych województw i uzyskano podobne wyniki. W województwie mazowieckim miasto stołeczne

można traktować jako metropolię, której oddziaływanie jest tak silne, że nie ma efektu grupowania się cen ziemi wokół małych miast.

PODSUMOWANIE

Na podstawie przeprowadzonych badań stwierdzono, że globalna statystyka i diagram Morana mogą być wykorzystane do klasyfikacji obiektów przestrzennych. Uzyskane podziały potwierdzono wykorzystując nieparametryczną analizę wariacji i badania szczegółowe. Podział na klastry potwierdził niezmienny w czasie i przestrzeni popyt na ziemię, przede wszystkim wokół dużych miast. Można przypuszczać, że wysoka cena ziemi rolniczej wokół dużych miast związana jest z rosnącymi potrzebami i przekształcaniem ziemi rolniczej w działki o charakterze budowlanym. Zaobserwowano również wpływ dużych miast na ceny gruntów rolniczych. Najsilniej było to obserwowane w województwie mazowieckim, jednak ta sytuacja występuje również w pozostałych województwach. Ze względu na ograniczenia objętościowe w pracy nie można było przedstawić szczegółowych analiz dla wszystkich województw oraz lat. Można je jednak znaleźć w pracy Pietrzykowskiego z 2019 r.

BIBLIOGRAFIA

- Anselin L. (1992) Space and Applied Econometrics. Special Issue: Regional Science and Urban Economics 22.
- Anselin L. (1995) Local Indicators of Spatial Association – LISA. *Geographical Analysis* 27(2), 93-115.
- Anselin L. (2001) *Spatial Econometrics*. Oxford: Basil Blackwell.
- Conover W. J. (1999) *Practical Nonparametric Statistics* (3rd ed.). New York: John Wiley and Sons.
- Dunn O. J. (1964) Multiple Comparisons using Rank Sums. *Technometrics*, (6), 241-252.
- Hanusz Z., Tarasińska J. (2014) On Multivariate Normality Tests using Skewness and Kurtosis. *Colloquium Biometricum*, 44, 139-148.
- Hubert L. J., Golledge R. G., Costanzo C. M., Gale N. (1985) Measuring Association between Spatially Defined Variables: an Alternative Procedure. *Geographical Analysis*, 17.
- Kopczewska K. (2020) *Przestrzenne metody ilościowe w R: statystyka, ekonometria, uczenie maszynowe, analiza danych*, Warszawa, CeDeWu.
- Kruskal W. H. (1952) A Nonparametric Test for the Several Sample Problem. *Annals of Mathematical Statistics*, 23, 525-540.
- Kruskal W. H., Wallis W. A. (1952) Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47, 583-621.
- Mardia K. V. (1970) Measures of Multivariate Skewness and Kurtosis with Applications. *Biometrika*, 57, 519-530.
- Moran P. (1948) The Interpretation of Statistical Maps. *Journal of the Statistical Royal Society, Series B*, 10, 243-251.

- Mantel N. (1967) The Detection of Disease Clustering and Generalized Regression Approach. *Cancer Research*, 27.
- Ord J. K., Getis A. (1995) Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, 27(4), 286-306.
- Ord J. K., Getis A. (2012) Local Spatial Heteroscedasticity (LOSH). *The Annals of Regional Science*, 48, 529-539.
- Pietrzykowski R. (2019) *Zróżnicowanie przestrzenne cen ziemi rolniczej w Polsce*. Warszawa, Wydawnictwo SGGW.
- Wilcoxon F. (1945) Individual Comparisons by Ranking Methods. *Biometrics*, 1(6), 80-83.

**CLASSIFICATION OF SPATIAL OBJECTS ON BASE THE GLOBAL
MORAN COEFFICIENT. ON THE EXAMPLE OF AGRICULTURAL
LAND PRICES**

Abstract: The aim of the work was to show the possibility of using the global Moran statistics for the classification of spatial objects on the example of agricultural land prices. The analyses were conducted at the level of communes (in NUTS 3 subregions of the voivodeship) in the years 2004-2012. Apart from the global Moran coefficient and the Moran diagram, nonparametric statistics (one-way ANOVA on ranks) were used in the work.

Keywords: spatial autocorrelation, agricultural land prices

JEL classification: C14, R3