

# PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

# RESEARCH PAPERS

of Wrocław University of Economics

Nr 427

**Taksonomia 27**

**Klasyfikacja i analiza danych –  
teoria i zastosowania**



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu  
Wrocław 2016

Redaktor Wydawnictwa: Agnieszka Flasińska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Tytuł dofinansowany ze środków Narodowego Banku Polskiego  
oraz ze środków Sekcji Klasyfikacji i Analizy Danych PTS

Informacje o naborze artykułów i zasadach recenzowania  
znajdują się na stronach internetowych  
[www.pracenaukowe.ue.wroc.pl](http://www.pracenaukowe.ue.wroc.pl)  
[www.wydawnictwo.ue.wroc.pl](http://www.wydawnictwo.ue.wroc.pl)

Publikacja udostępniona na licencji Creative Commons  
Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 3.0 Polska  
(CC BY-NC-ND 3.0 PL)



© Copyright by Uniwersytet Ekonomiczny we Wrocławiu  
Wrocław 2016

**ISSN 1899-3192** (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)  
**e-ISSN 2392-0041**  
**ISSN 1505-9332** (Taksonomia)

Wersja pierwotna: publikacja drukowana

Zamówienia na opublikowane prace należy składać na adres:  
Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu  
ul. Komandorska 118/120, 53-345 Wrocław  
tel./fax 71 36 80 602; e-mail:[econbook@ue.wroc.pl](mailto:econbook@ue.wroc.pl)  
[www.ksiegarnia.ue.wroc.pl](http://www.ksiegarnia.ue.wroc.pl)

Druk i oprawa: TOTEM

## Spis treści

<b>Wstęp</b> .....	9
<b>Beata Bal-Domańska:</b> Propozycja procedury oceny zrównoważonego rozwoju w układzie <i>presja – stan – reakcja</i> w ujęciu przestrzennym / Proposal of the assessment of poviats sustainable development in the pressure – state – response system in spatial terms.....	11
<b>Tomasz Bartłomowicz:</b> Pomiar preferencji konsumentów z wykorzystaniem metody <i>Analytic Hierarchy Process</i> / Analytic Hierarchy Process as a method of measurement of consumers’ preferences.....	20
<b>Maciej Beręsewicz, Marcin Szymkowiak:</b> Analiza skupień wybranych lokalnych rynków nieruchomości w Polsce z wykorzystaniem internetowych źródeł danych / Cluster analysis of selected local real estate markets in Poland based on Internet data sources.....	30
<b>Beata Bieszk-Stolorz:</b> Wybrane modele przeciętnego efektu oddziaływania w analizie procesu wychodzenia z bezrobocia / Chosen average treatment effect models in the analysis of unemployment exit process.....	40
<b>Justyna Brzezińska:</b> Modele IRT i modele Rascha w badaniach testowych / IRT and Rasch models in test measurement.....	49
<b>Mariola Chrzanowska, Nina Drejerska:</b> Geograficznie ważona regresja jako narzędzie analizy poziomu rozwoju społeczno-gospodarczego na przykładzie regionów Unii Europejskiej / Geographically weighted regression as a tool of analysis of socio-economic development level of regions in the European Union.....	58
<b>Sabina Denkowska:</b> Zastosowanie analizy wrażliwości do oceny wpływu nieobserwowanej zmiennej w <i>Propensity Score Matching</i> / The application of sensitivity analysis in assessing the impact of an unobserved confounder in Propensity Score Matching.....	66
<b>Adam Depta:</b> Zastosowanie analizy czynnikowej do wyodrębnienia aspektów zdrowia wpływających na jakość życia osób jaskających się / The application of factor analysis to the identification of the health aspects affecting the quality of life of stuttering people.....	76
<b>Mariusz Doszyń, Sebastian Gnat:</b> Taksonomiczno-ekonometryczna procedura wyceny nieruchomości dla różnych miar porządkowania / Taxonomic and econometric method of real estate valuation for various classification measures.....	84

<b>Marta Dziechciarz-Duda, Anna Król:</b> Segmentacja konsumentów smartfonów na podstawie preferencji wyrażonych / Segmentation of smartphones' consumers on the basis of stated preferences .....	94
<b>Ewa Genge:</b> Zmienne towarzyszące w ukrytym modelu Markowa – analiza oszczędności polskich gospodarstw domowych / Latent Markov model with covariates – Polish households' saving behaviour .....	103
<b>Joanna Górna, Karolina Górna:</b> Modelowanie wzrostu gospodarczego z wykorzystaniem narzędzi ekonometrii przestrzennej / Economic growth modelling with the application of spatial econometrics tools .....	112
<b>Alicja Grześkowiak:</b> Wielowymiarowa analiza kompetencji zawodowych według grup wieku ludności / Multivariate analysis of professional competencies with respect to the age groups of the population .....	122
<b>Agnieszka Kozera, Feliks Wysocki:</b> Problem ustalania współrzędnych obiektów modelowych w metodach porządkowania liniowego obiektów / The problem of determining the coordinates of model objects in object linear ordering methods .....	131
<b>Mariusz Kubus:</b> Lokalna ocena mocy dyskryminacyjnej zmiennych / Local evaluation of a discrimination power of the variables.....	143
<b>Paweł Lula, Katarzyna Wójcik, Janusz Tuchowski:</b> Analiza wydźwięku polskojęzycznych opinii konsumenckich ukierunkowanych na cechy produktu / Feature-based sentiment analysis of opinions in Polish.....	153
<b>Aleksandra Łuczak, Agnieszka Kozera, Feliks Wysocki:</b> Ocena sytuacji finansowej jednostek samorządu terytorialnego z wykorzystaniem rozmytych metod klasyfikacji i programu R / Assessment of financial condition of local government units with the use of fuzzy classification methods and program R .....	165
<b>Dorota Rozmus:</b> Badanie stabilności taksonomicznej czynnikowej metody odległości probabilistycznej / Stability of the factor probability distance clustering method .....	176
<b>Adam Sagan, Aneta Rybicka, Justyna Brzezińska:</b> <i>Conjoint analysis</i> oparta na modelach IRT w zagadnieniu optymalizacji produktów bankowych / An IRT-approach for conjoint analysis for banking products preferences.....	184
<b>Michał Stachura:</b> O szacowaniu centrum populacji określonego obszaru na przykładzie Polski / On estimating centre of population of a given territory. Poland's case .....	195
<b>Michał Stachura, Barbara Wodecka:</b> Wybrane aspekty i zastosowania modeli zdarzeń ekstremalnych / Selected facets and application of models of extremal events .....	205
<b>Iwona Staniec, Jan Żółtowski:</b> Wykorzystanie analizy log-liniowej do wyboru czynników determinujących współpracę w przedsiębiorczości	

---

technologicznej / Use of log-linear analysis for the selection determinants of cooperation in technological entrepreneurship.....	215
<b>Marcin Szymkowiak, Wojciech Roszka:</b> Potencjał gospodarczy gmin aglomeracji poznańskiej w ujęciu taksonomicznym / The economic potential of municipalities of the Poznań agglomeration in the light of taxonomy analysis.....	224
<b>Lucyna Wojcieszka:</b> Zastosowanie modeli klas ukrytych w badaniu opinii respondentów na temat roli państwa w gospodarce / Implementation of latent class models in the respondents' survey on the role of the country in economy.....	234

## **Wstęp**

W dniach 14–16 września 2015 r. w Hotelu Novotel Gdańsk Marina w Gdańsku odbyła się XXIV Konferencja Naukowa Sekcji Klasyfikacji i Analizy Danych PTS (XXIX Konferencja Taksonomiczna) „Klasyfikacja i analiza danych – teoria i zastosowania”, zorganizowana przez Sekcję Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego oraz Katedrę Statystyki Wydziału Zarządzania Uniwersytetu Gdańskiego.

W trakcie dwóch sesji plenarnych oraz 13 sesji równoległych wygłoszono 58 referatów poświęconych aspektom teoretycznym i aplikacyjnym zagadnienia klasyfikacji i analizy danych. Odbyła się również sesja plakatowa, na której zaprezentowano 14 plakatów.

Teksty 24 recenzowanych artykułów naukowych stanowią zawartość prezentowanej publikacji z serii Taksonomia nr 27. Teksty 25 recenzowanych artykułów naukowych znajdują się w Taksonomii nr 26.

*Krzysztof Jajuga, Marek Walesiak*

**Mariola Chrzanowska, Nina Drejerska**

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie  
e-mail: {mariola\_chrzanowska; nina\_drejerska}@sggw.pl

---

**GEOGRAFICZNIE WAŻONA REGRESJA  
JAKO NARZĘDZIE ANALIZY POZIOMU ROZWOJU  
SPOŁECZNO-GOSPODARCZEGO  
NA PRZYKŁADZIE REGIONÓW UNII EUROPEJSKIEJ**

---

**GEOGRAPHICALLY WEIGHED REGRESSION  
AS A TOOL OF ANALYSIS OF SOCIO-ECONOMIC  
DEVELOPMENT LEVEL OF REGIONS  
IN THE EUROPEAN UNION**

---

DOI: 10.15611/pn.2016.427.06

**Streszczenie:** Łączenie wątków związanych z teoriami wzrostu endogenicznego oraz nową geografią ekonomiczną w badaniach nad wzrostem społeczno-gospodarczym doprowadziło do rozwoju *New Economic Geography and Growth* (NEGG). Metoda regresji ważonej geograficznie (GWR) jest w takich ramach właściwym narzędziem analizy, ponieważ uwzględniając czynniki rozważane w teoriach rozwoju endogenicznego (np. kapitał ludzki czy kapitał intelektualny), pozwala jednocześnie wziąć pod uwagę położenie badanych obiektów w przestrzeni geograficznej. W opracowaniu poddano analizie poziom rozwoju społeczno-gospodarczego regionów w Unii Europejskiej z wykorzystaniem geograficznie ważonej regresji. Jako zmienną zależną wybrano PKB *per capita* według parytetu siły nabywczej, a jako zmienne niezależne charakterystyki kapitału ludzkiego i kapitału intelektualnego. Zaprezentowane wyniki badań wskazują na zasadność stosowania modeli uwzględniających przestrzenną niestacjonarność, w tym np. GWR.

**Słowa kluczowe:** geograficznie ważona regresja, rozwój regionalny, przestrzenna niestacjonarność.

**Summary:** Combination of issues connected with theories of endogenous growth with the new economic geography in research on socio-economic growth led to the development of the new economic geography and growth (NEGG). The method of geographically weighted regression is a proper analytical instrument within this approach, as it concerns factors coming from theories of endogenous growth (e.g. human or intellectual capital) as well as location of investigated objects. The paper presents analysis of a level of socio-economic development of regions in the European Union with the use of the geographically weighted regression. Gross Domestic Product in PPS *per capita* was used as a dependent variable whereas characteristics of human and intellectual capital were applied as independent variables. Presented research results indicate that it is reasonable to use models referring to spatial nonstationarity, as for example the geographically weighted regression.

**Keywords:** geographically weighted regression, regional development, spatial nonstationarity.

## 1. Wstęp

Porównania poziomu rozwoju społeczno-gospodarczego w Unii Europejskiej dokonywane są często pomiędzy poszczególnymi państwami członkowskimi. W takim ujęciu trudno jest porównywać małe państwa, takie jak np. Malta (425 tys. mieszkańców) czy Luksemburg (550 tys. mieszkańców) z największymi, takimi jak np. Republika Federalna Niemiec (prawie 81 milionów mieszkańców). Dlatego też analizowanie danych na poziomie regionów (NUTS 2) UE pozwala na lepszą charakterystykę sytuacji i procesów społeczno-gospodarczych [Eurostat 2015, s. 8]. Regionalne podejście zastosowali np. J. Crespo-Cuaresma, N. Foster i R. Stehrer [2011] czy S. Dall'Erba, M. Percoco i G. Piras [2008], jakkolwiek przedstawiane są również modele łączące czynniki analizowane na poziomie krajowym (np. poziom inflacji) i regionalnym (np. poziom kapitału ludzkiego) przez R. Capello i U. Fratesi [2012].

Badania nad rozwojem regionów i polaryzacją przestrzeni ekonomicznej są jednym z głównych trendów badawczych w nowej geografii ekonomicznej. Polaryzację tej przestrzeni, jej istotę, efekty i przyczyny interpretuje się na podstawie wielu teorii, opartych głównie na ścieżce rozwoju nierównomiernego. Włączenie przestrzeni do teorii wzrostu przez połączenie teorii wzrostu endogenicznego i nowej geografii ekonomicznej prowadzi do prezentacji w literaturze *New Economic Geography and Growth* (NEGG) [Bond-Smith, McCann 2014]. Modelowanie wzrostu według tego podejścia zakłada uwzględnienie m.in. takich czynników przestrzennych, jak lokalizacja. Jednocześnie nadal w modelowaniu wzrostu społeczno-gospodarczego brane są pod uwagę czynniki przywoływane w ramach neoklasycznych teorii wzrostu, np. wyposażenie w infrastrukturę [OECD 2009].

Celem opracowania jest określenie poziomu rozwoju społeczno-gospodarczego regionów (NUTS 2) Unii Europejskiej z zastosowaniem geograficznie ważonej regresji (GWR), ze szczególnym uwzględnieniem regionów (województw) Polski. Uwzględnienie zależności przestrzennych, możliwe przy wykorzystaniu GWR, umożliwia analizę poziomu rozwoju w kontekście lokalnym przy wykorzystaniu informacji globalnych.

W pracy wykorzystano dane pochodzące z Eurostatu za 2011 r.<sup>1</sup> Interpretacji wyników z wykorzystaniem mapy dokonano dla regionów Polski. Zaprezentowane wyniki mają wartość poznawczą, natomiast mogą też być inspiracją dla dalszych badań regionalnych uwzględniających zależności przestrzenne.

---

<sup>1</sup> Najbardziej aktualne dane o poziomie PKB na mieszkańca w układzie regionalnym Unii Europejskiej były dostępne dla 2011 r.



## 2. Geograficznie ważona regresja

Chociaż przestrzenne zróżnicowanie regionów jest oczywiste i empirycznie udowodnione, to świadomość istnienia zależności przestrzennych nie zawsze skutkuje stosowaniem odpowiednich (tj. uwzględniających te zależności) metod statystycznych w analizie rozwoju regionalnego. Budowę modelu z uwzględnieniem zależności przestrzennych umożliwiają m.in. regresja ważona geograficznie (GWR). Ta technika eksploracyjna jest rozbudowaną wersją modelu regresji, ponieważ przedstawia lokalne własności, które są niewidoczne w podejściu standardowym.

Metoda regresji z wagowaniem geograficznym „pozwała na uchwycenie dla każdej obserwacji zmienności współczynników regresji w przestrzeni” [Suchecki 2010, s. 264]. Estymacja parametrów lokalnych modeli regresji dokonywana jest z uwzględnianiem charakterystyk (zmiennych objaśniających) pochodzących przede wszystkim z obserwacji sąsiednich, z malejącym wraz z odległością znaczeniem kolejnych obserwacji.

W klasycznym podejściu (tzn. budowie jednego, globalnego modelu) zakłada się stacjonarność (homogeniczność) przestrzenną badanego procesu. Innymi słowy, w klasycznej regresji w każdej części analizowanego obszaru badany proces podlega identycznym uwarunkowaniom, a określone oddziaływanie zmiennych diagnostycznych przynosi takie same rezultaty. Związek pomiędzy zmiennymi objaśniającymi a zmienną objaśnianą jest stały w przestrzeni i model może być estymowany MNK. Co więcej, przestrzenna heteroskedastyczność nie wpływa na istotność oszacowanych parametrów, co w konsekwencji może doprowadzić do błędnej interpretacji.

Model regresji ważonej geograficznie (GWR) to jeden z modeli heterogeniczności przestrzennej, które uwzględniają brak stabilności przestrzennej parametrów strukturalnych. Innymi słowy GWR bazuje na założeniu, że (dopasowane do wszystkich danych) oszacowane wartości współczynników modelu globalnego mogą nie uwzględniać w odpowiedni sposób różnic lokalnych [Bivand, Pebesma, Gomez-Rubio 2008]. Zgodnie z tym założeniem regresja ważona geograficznie polega na budowaniu lokalnych modeli regresji z wykorzystaniem macierzy sąsiedztwa. Innymi słowy, parametry modelu szacowane są osobno dla każdego obiektu w przestrzeni, dla którego znane są wartości zmiennych objaśnianej i objaśniających [Suchecki 2010]. W tym przypadku estymowane parametry równania są związane z lokalizacją obiektów. Model regresji ważonej geograficznie dla  $i$ -tej obserwacji można opisać wzorem:

$$y_i = \beta_0(x_0, y_0) + \beta_1(x_1, y_1)x_{1i} + \beta_2(x_2, y_2)x_{2i} + \dots + \beta_k(x_i, y_i)x_{ik} + \varepsilon_i, \quad (1)$$

gdzie:  $y_i$  – wartość zmiennej niezależnej dla  $i$ -tego obiektu;  $x_{ik}$  – wartość  $k$ -tej zmiennej zależnej dla  $i$ -tego obiektu;  $(x_i, y_i)$  – współrzędne geograficzne  $i$ -tego obiektu dla  $i = 1, \dots, n$ ;  $\beta_k(x_i, y_i)$  – realizacja globalnej funkcji  $\beta_k(x, y)$  w punkcie  $i$ .

Wartości parametrów uzyskiwane są za pomocą ważonej metody najmniejszych kwadratów:

$$\sum_{i=1}^n w_{i(i)} (y_i - (b_0(x_0, y_0) + b_k(x_i, y_i) x_{ik}))^2 \rightarrow \min. \quad (2)$$

Technikę estymacji parametrów GWR dokładnie opisali A.S. Fotheringham, C. Brunsdon i M. Charlton [2002]. Aby możliwe było ustalenie wartości lokalnych parametrów modelu, przyjmuje się założenie, iż dla danej lokalizacji bardziej istotne są obserwacje z obszarów lub punktów sąsiednich niż obserwacje bardziej oddalone. W tym celu różnicuje się informacje, tworząc macierz wag  $\mathbf{W}$ . Elementy diagonalne tej macierzy to odległości między  $i$ -tą lokalizacją a wszystkimi pozostałymi lokalizacjami z bazy. Innymi słowy, główna przekątna macierzy  $\mathbf{W}$  zawiera informację o wagach geograficznych odnoszących się do otoczenia danej lokalizacji [Suchecki 2010]. Kolejne elementy macierzy wyznaczone są następująco<sup>2</sup>:

$$w_{ij} = e^{\left(\frac{-d_{ij}^2}{h^2}\right)}, \quad (3)$$

gdzie:  $w_{ij}$  – element macierzy wag,  $d_{ij}$  – odległość między  $i$ -tym a  $j$ -tym obszarem;  $h$  – parametr opóźnienia (*bandwidth*).

Parametr  $h$  (okno) określa liczbę obserwacji wykorzystanych do estymacji współczynników modelu lokalnego. Jeżeli wartość  $h$  dąży do nieskończoności, to wagi dążą do jedności. Oceny parametrów uzyskane wówczas w wyniku zastosowania GWR są ekwiwalentne ocenom MNK. Z kolei gdy wartość  $h$  maleje, oceny oszacowanych parametrów zależą coraz bardziej od obserwacji zlokalizowanych najbliżej, a ich wariancja rośnie. „Okno”  $h$  informuje zatem o stopniu wygładzenia danych [Suchecki 2010]. Unikalny zestaw oszacowanych parametrów dla każdej obserwacji wyznacza się na podstawie wzoru:

$$\beta(x_i, y_i) = (\mathbf{X}^T \mathbf{W}_{(i)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_{(i)} \mathbf{y}. \quad (4)$$

Dla wyznaczonych parametrów można analizować miary dopasowania do danych empirycznych. Wysokie zróżnicowanie oszacowanych wartości ocen parametrów wskazuje na przestrzenną heterogeniczność zjawiska.

Do oceny jakości zbudowanego modelu stosuje się zwykle (wykorzystywane również w analizie regresji liniowej) współczynnik determinacji oraz kryterium Akaike (AIC). Weryfikacji hipotezy o istotności statystycznej różnic ocen parametrów dokonuje się na podstawie wybranych testów parametrycznych (np. test

<sup>2</sup> Inne sposoby wyznaczania wag, wraz z dyskusją znaleźć można w pracy [Brundson, Fortheringham, Charlton 1998].

ANOVA GWR). Są one opisane w pracy m.in. w pracy [Brundston, Fortheringham, Charlton 1999].

Metoda regresji ważonej geograficznie jest stosowana w literaturze np. do badania poziomu rozwoju społeczno-gospodarczego wyrażonego wartością PKB na mieszkańca [Fabian 2014]. GWR stosowano również do przestrzennej analizy migracji zagranicznych w Polsce [Lewandowska-Gwarda, 2014] czy rozwoju turystyki w Chinach [Yang, Fik 2014].

### 3. Opis badania

W pierwszym etapie badań oszacowano model globalny. Jako zmienną zależną opisującą poziom rozwoju społeczno-gospodarczego regionów wybrano PKB *per capita* według parytetu siły nabywczej. Pod uwagę wzięto czynniki wzrostu społeczno-gospodarczego wykorzystywane w modelach wzrostu regionalnego OECD [2009], opisujące: infrastrukturę transportową, poziom wykształcenia mieszkańców, wydatki na badania i rozwój, zatrudnienie w działalności badawczo-rozwojowej oraz w sektorach wysokiej technologii, aplikacje do Europejskiego Urzędu Patentowego oraz ogólny wskaźnik zatrudnienia. Spośród zaproponowanego zestawu zmiennych metodą regresji krokowej wstecz wybrano trzy zmienne:

- odsetek osób z wykształceniem wyższym w grupie wiekowej 25–64 lata ( $X_1$ ),
- odsetek zatrudnionych w sektorach wysokich technologii w zatrudnionych ogółem ( $X_2$ ),
- liczbę aplikacji do Europejskiego Urzędu Patentowego na milion mieszkańców ( $X_3$ ).

Wybrane informacje o modelu zaprezentowano w tab. 1 (z analiz wyłączono regiony, dla których informacje były niepełne lub niewiarygodne; są to regiony leżące poza kontynentem – Guadelupa, Martynika, Gujana).

Analizując informacje przedstawione w tab. 1, można zauważyć, że zmienność PKB jest tłumaczona w 50% za pomocą wybranych zmiennych diagnostycznych. Wyznaczone parametry modelu globalnego są statystycznie istotne. Znaki stojące

**Tabela 1.** Podstawowe informacje o modelu regresji wielokrotnej krokowej wstecznej

$N = 273$	$R^2 = 0,50$ Skorygowany $R^2 = 0,44$ Kryterium Informacyjne Akaike AIC = 5584,65 Test White'a (Statystyka $LM = 97,036$ ; wartość $p = 0,00$ )			
	współczynnik	błąd std	Statystyka $t$	wartość $p$
Stała	9102,71	1289,73	7,55	0,00
$X_1$	329,41	47,50	6,94	0,00
$X_2$	60,51	17,27	3,50	0,00
$X_3$	38,73	4,27	9,07	0,00

Źródło: opracowanie własne na podstawie danych Eurostatu.

przy oszacowanych parametrach potwierdzają dodatni wpływ zmiennych diagnostycznych na badane zjawisko. Wyniki testu White'a potwierdzają występowanie heteroskedastyczności reszt, a przestrzenny charakter tego zjawiska potwierdziła wyznaczona dla reszt modelu Statystyka Morana ( $I = 0,26$ ). A zatem istnieją przesłanki do zastosowania GWR w dalszych badaniach. Szczegółowe informacje dla GWR przedstawiono w tab. 2.

**Tabela 2.** Podstawowe informacje o modelu GWR

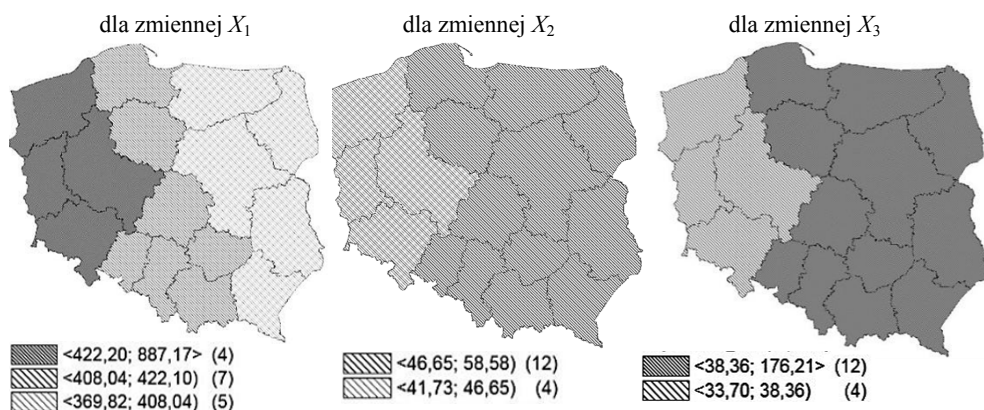
Charakterystyki opisowe	Współczynniki			
	Stała	$X_1$	$X_2$	$X_3$
Średnia	9137,79	388,74	54,10	33,43
Odch. standardowe	2531,85	84,94	32,24	15,04
Kwartył pierwszy (Q1)	7581,84	369,82	41,73	30,15
Mediana (Q2)	9531,31	408,04	46,65	33,70
Kwartył trzeci (Q3)	10598,94	422,20	58,58	38,36
Min.	1502,85	-375,71	28,72	-119,48
Max.	19488,42	511,47	467,52	56,94
Rozstęp	17985,57	887,18	438,60	176,42
$R^2 = 0,53$ Skorygowany $R^2 = 0,50$ Kryterium Informacyjne Akaike AIC = 5559,09.				

Źródło: opracowanie własne na podstawie danych Eurostatu.

Zgodnie z kryterium AIC model GWR jest lepszy od modelu globalnego. Dla regresji geograficznie ważonej wartość tego kryterium jest niższa. Na lepszą jakość GWR wskazuje również nieznacznie wyższa wartość współczynnika determinacji. Dodatkowo, na korzyść GWR „przemawia” duże zróżnicowanie ocen wyznaczonych lokalnie parametrów – najniższa różnica pomiędzy wartością minimalną a maksymalną to 176. Przestrzenne zróżnicowanie parametrów modelu potwierdzono za pomocą testu ANOVA. Hipoteza zerowa w tym teście informuje, że model GWR generuje wyniki na tym samym poziomie co model globalny. W prezentowanym badaniu wyznaczona wartość  $F = 3,59$  (wartość  $p < 0,05$ ) wskazuje na istotną różnicę pomiędzy modelem globalnym i modelem GWR. Innymi słowy, za pomocą modelu GWR otrzymano lepsze oceny parametrów modelu.

Estymacja modelu za pomocą GWR umożliwiła analizę zróżnicowania pozyskanych wyników lokalnie dla poszczególnych regionów. Oszacowane estymatory parametrów zostały podzielone na cztery grupy na podstawie kwartyli. Wyniki badań dla regionów Polski zaprezentowano na rys. 1.

Największe zróżnicowanie zaobserwowano dla zmiennej  $X_1$ . W tym przypadku wartości z przedziału pomiędzy Q1 a Q2 zaobserwowano głównie dla grupy województw: warmińsko-mazurskiego, podlaskiego, mazowieckiego, lubelskiego i podkarpackiego. Wielkości z zakresu Q2 pomiędzy Q3 zaobserwowano głównie dla



Rys. 1. Wartości parametrów lokalnej regresji geograficznie ważonej regionów Polski

Źródło: opracowanie własne danych Eurostatu.

pasa regionów Polski centralnej. Najwyższa wartość estymatorów parametrów (powyżej Q3) cechowała województwa zachodnie. Analizując wyniki dla zmiennych  $X_2$  i  $X_3$  można zauważyć wyraźny podział regionów na grupę A i B. Co ciekawe, w obu przypadkach nie zanotowano grupy obiektów o wartościach parametrów poniżej Q1.

#### 4. Podsumowanie

Głównym celem pracy było określenie poziomu rozwoju społeczno-gospodarczego regionów (NUTS 2) Unii Europejskiej z zastosowaniem geograficznie ważonej regresji (GWR), ze szczególnym uwzględnieniem regionów (województw) Polski. Regionalny rozwój jest pojęciem trudnym do modelowania i w tym kontekście każda próba takiej analizy jest wskazana.

Rezultaty otrzymane dla modelu globalnego nie były satysfakcjonujące, a duża heterogeniczność sugerowała wykorzystanie metody uwzględniającej przestrzenne zróżnicowanie. Przestrzenną heteroskedastyczność potwierdziła również statystycznie istotna dodatnia globalna statystyka Morana. Statystyki opisowe obliczone dla estymatorów parametrów oszacowanych dla modelu GWR potwierdziły duże zróżnicowanie zmiennych objaśniających. Dodatkowo wyznaczone miary średnie były zbliżone do wartości oszacowań dla modelu globalnego, co przy dużym zróżnicowaniu wyraźnie przemawia na korzyść modelu GWR.

Statystyczna weryfikacja testem ANOVA modelu globalnego z modelem GWR pozwoliła na sformułowanie wniosku o lepszym dopasowaniu modelu GWR do badanego zjawiska. Zastosowany model GWR trafnie przewidział pozycję regionów Polski na tle innych regionów UE. Estymatory parametrów dla regionów Pol-

ski kształtowały się zwykle na wyższym poziomie niż Q1, przy czym można było zaobserwować wyraźnie wyższe wartości dla województw zachodnich w porównaniu z resztą kraju.

## Literatura

- Bivand R., Pebesma E.J., Gomez-Rubio V., 2008, *Applied Spatial Data analysis in R*, Springer, New York.
- Bond-Smith S., McCann P., 2014, *Incorporating space in the theory of endogenous growth*, [w:] M.M. Fischer, P. Nijkamp (red.), *Handbook of Regional Science*, Springer Reference.
- Brundson C., Fortheringham A.S., Charlton M., 1998, *Geographically weighted regression: A method for exploring spatial nonstationarity*, *Journal of the Royal Statistical Society. Series D (The Statistician)*, vol. 47, no. 3, s. 431–443.
- Brundson C., Fortheringham A.S., Charlton M., 1999, *Some notes on parametric significance tests for geographically weighted regression*, *Journal of Regional Science*, vol. 39, no. 3, s. 497–524.
- Capello R., Fratesi U., 2012, *Modelling regional growth: An advanced MASST model*, *Spatial Economic Analysis*, vol. 7, no. 3, s. 293–318.
- Crespo-Cuaresma J., Foster N., Stehrer R., 2011, *Determinants of regional economic growth by quantile*, *Regional Studies*, vol. 45, no. 6, s. 809–826.
- Dall’Erba S., Percoco M., Piras G., 2008, *The European regional growth process revisited*, *Spatial Economic Analysis*, vol. 3, no. 1, s. 7–25.
- Eurostat, 2015, *Eurostat Regional Yearbook 2015*, Publications Office of the European Union, Luxembourg.
- Fabian Z., 2014, *Method of the geographically weighted regression and example for its application*, *Regional Statistics*, vol. 4, no. 1, s. 61–75.
- Fortheringham A.S., Brundson C., Charlton M., 2002, *Geographically Weighted Regression the Analysis of Spatially Varying Relationships*, John Wiley & Sons, Hoboken, NJ.
- Lewandowska-Gwarda K., 2014, *Spatial analysis of foreign migration in Poland in 2012 using geographically weighted regression*, *Comparative Economic Research*, vol. 17, no. 4, s. 137–154.
- OECD, 2009, *How Regions Grow. Trends and Analysis*, OECD Publishing, Paris.
- Sucheckı B., 2010, *Ekonometria przestrzenna, Metody i modele analizy danych przestrzennych*, C.H. Beck, Warszawa.
- Yang Y., Fik T., 2014, *Spatial effects in regional tourism growth*, *Annals of Tourism Research*, vol. 46, s. 144–162.