

Michał Bernard Pietrzak

**PROBLEM ZMIENNEJ JEDNOSTKI ODNIESIENIA
W PRZESTRZENNYCH BADANIACH EKONOMICZNYCH**

Polskie Towarzystwo Ekonomiczne
Oddział w Toruniu

Toruń 2018

Recenzenci:

prof. zw. dr hab. Jan Zawadzki
dr hab. Elżbieta Szulc, prof. UMK

Skład, łamanie i projekt okładki:

Ilona Pietryka

© Copyright by Polskie Towarzystwo Ekonomiczne Oddział w Toruniu

ISBN 978-83-62049-32-5

DOI: 10.24136/eep.mon.2018.1

Polskie Towarzystwo Ekonomiczne Oddział w Toruniu
e-mail: pte@umk.pl
www.ptetorun.pl

ul. Kopernika 21
87–100 Toruń

Praca powstała w wyniku realizacji projektu badawczego nr rej. 2015/17/B/HS4/01004 finansowanego ze środków Narodowego Centrum Nauki.

Spis treści

Wstęp	7
Rozdział 1. Agregacja danych w przestrzennych badaniach ekonomicznych: aspekty teoretyczne	13
1.1. Podstawowe zagadnienia dotyczące procesu agregacji danych	13
1.1.1. Agregacja danych a problem zmiennej jednostki odniesienia	13
1.1.2. Agregacja danych w badaniach ekonomicznych	16
1.2. Systematyka agregacji danych	21
1.2.1. Typy agregacji danych	21
1.2.2. System agregacji danych	26
1.2.3. Specyfika agregacji danych przestrzennych	29
Rozdział 2. Poprawność przestrzennych badań ekonomicznych	37
2.1. Etapy przestrzennych badań ekonomicznych	37
2.1.1. Etapy agregacji danych przestrzennych	37
2.1.2. Etap przestrzennej analizy ekonomicznej	42
2.2. Potencjalne błędy badawcze w przestrzennych badaniach ekonomicznych	48
2.2.1. Błąd identyfikacji jednorodnego zbioru układów	48
2.2.2. Błąd wyznaczenia obszaru wniosków	57
2.3. Identyfikacja mikrozależności i makrozależności w badaniach ekonomicznych	60
2.3.1. Mikrozależności a makrozależności ekonomiczne	60
2.3.2. Analiza mikrozależności oraz makrozależności przestrzennych	62
Rozdział 3. Redefinicje problemu sposobu podziału przestrzeni oraz problemu skali	75
3.1. Geneza powstania problemu zmiennej jednostki odniesienia	75
3.2. Problem zmiennej jednostki odniesienia a nowy paradygmat badań naukowych	79
3.2.1. Nowy paradygmat badań naukowych	79

3.2.2. Problem zmienności otrzymywanych rezultatów badań naukowych	83
3.2.3. Problem modyfikowalności jednostek terytorialnych	88
3.2.4. Wyznaczanie granic układu jednostek terytorialnych a problem zmiennej jednostki odniesienia	90
3.2.5. Jednorodność przyczynowa danych przestrzennych a wyniki badań	94
3.2.6. Potencjalne scenariusze badań nad problemem zmiennej jednostki odniesienia	97
3.2.7. Metodyka badań naukowych w ramach nowego paradygmatu nauk empirycznych	99
3.3. Rewizja problemu zmiennej jednostki odniesienia	102
3.3.1. Redefinicja problemu sposobu podziału przestrzeni	102
3.3.2. Redefinicja problemu skali	104
Rozdział 4. Podział przestrzeni dla badań dotyczących rozwoju rolnictwa w Polsce	107
4.1. Rozwiązanie problemu sposobu podziału przestrzeni na poziomie makroregionów	107
4.1.1. Wyznaczenie jednorodnego zbioru układów	107
4.1.2. Wyznaczenie granic układu makroregionów rolniczych	113
4.2. Relacja między mikroparametrami a makroparametrem na przykładzie badania makroregionów rolniczych SGM	121
4.2.1. Identyfikacja mikrozależności i makrozależności przestrzennych dotyczących zużycia nawozów mineralnych w rolnictwie	121
4.2.2. Wyznaczenie makrozależności przestrzennej na podstawie mikrozależności przestrzennych	131
Rozdział 5. Problem skali: identyfikacja na podstawie generowanych danych przestrzennych wyrażonych w wartościach średnich	141
5.1. Procedura symulacyjna w świetle zaproponowanej redefinicji problemu skali	141
5.2. Analiza symulacyjna dla szumów przestrzennych	143
5.3. Identyfikacja problemu skali w przypadku występowania trendu przestrzennego	153
5.4. Analiza wpływu autokorelacji przestrzennej na problem skali	162
5.4.1. Identyfikacja struktury przestrzennej	162
5.4.2. Wybór rozkładu losowego oraz systemu wag w procedurze symulacyjnej	165
5.4.3. Wpływ siły autokorelacji przestrzennej na poziom zmian wynikających z problemu skali	177
Rozdział 6. Symulacyjna analiza problemu skali na podstawie ważonych danych przestrzennych	183
6.1. Identyfikacja problemu skali dla szumów przestrzennych	183
6.2. Badanie problemu skali w warunkach występowania niejednorodności systematycznej	193
6.3. Analiza symulacyjna dla procesów o własności autokorelacji przestrzennej	206

Zakończenie	215
Literatura	221
Spis tabel	233
Spis rysunków	237

Wstęp

Praca dotyczy problemu zmiennej jednostki odniesienia¹. W ramach problemu zmiennej jednostki odniesienia rozpatrywane są osobno dwa zagadnienia, problem sposobu podziału przestrzeni oraz problem skali (poziomu agregacji)². Istota obydwu problemów określona została w literaturze w kontekście możliwości uzyskania odmiennych wyników badań na skutek zmiany układu jednostek terytorialnych lub w wyniku wykonania procesu agregacji danych.

Problem sposobu podziału przestrzeni opisany jest w literaturze jako problem związany z otrzymywaniem różnych wyników na podstawie analizy danych przestrzennych w zależności od wyboru układu jednostek terytorialnych, jednak w ramach tego samego poziomu agregacji (Openshaw, Taylor, 1979, s. 128; Openshaw, 1984b, s. 8). Natomiast problem skali określany jest jako problem związany z uzyskaniem odmiennych wyników badań pod wpływem zmiany poziomu agregacji danych przestrzennych (Openshaw, Taylor, 1979, s. 128; Openshaw, 1984b, s. 8).

Celem pracy jest opracowanie nowego podejścia metodologicznego do identyfikacji problemu zmiennej jednostki odniesienia, w tym problemu podziału przestrzeni oraz problemu skali w przestrzennych badaniach ekonomicznych. Właściwe zrozumienie tych problemów i ich uwzględnienie w prowadzonych przestrzennych badaniach ekonomicznych ma istotny wpływ na otrzymywane wyniki.

¹ Termin „problem zmiennej jednostki odniesienia” przyjęty został z pracy Suhecka (2014, s. 51–52, 56–64). Stanowi on tłumaczenie z języka angielskiego wyrażenia „Modifiable Areal Unit Problem” (MAUP).

² Również terminy „problem sposobu podziału przestrzeni” oraz „problem skali” („problem poziomu agregacji”) przyjęte zostały z pracy Suhecka (2014, s. 56–64) i stanowią one tłumaczenie z wyrażen „Aggregation Problem”, „Zonation Problem” oraz „Scale Problem” (Openshaw, Taylor 1979, s. 128, Openshaw, 1984b, s. 8).

Przeprowadzone w pracy rozważania koncentrują się wokół następujących tez:

1. Problem podziału przestrzeni oraz problem skali wymagają redefinicji, uwzględniających kwestię jednorodności przyczynowej danych przestrzennych.
2. Ustalenie jednorodnego zbioru układów w ramach prowadzonych przestrzennych badań ekonomicznych pozwala na poprawną identyfikację własności jednorodności przyczynowej danych przestrzennych.
3. Błąd identyfikacji jednorodnego zbioru układów oraz błąd wyznaczania obszaru wniosków prowadzą do niepoprawnych wniosków formułowanych na podstawie wyników analiz przestrzennych.
4. Identyfikacja makrozależności przestrzennych często obciążona jest niebezpieczeństwem popełnienia błędu wyznaczenia obszaru wniosków.
5. Zastosowanie *zoning system* lub *grouping system* do losowego wyznaczania granic układów terytorialnych nie gwarantuje, że odniesione do nich dane przestrzenne charakteryzują się własnością jednorodności przyczynowej.
6. Identyfikacja problemu zmiennej jednostki odniesienia na podstawie wyznaczanych w sposób losowy układów jednostek terytorialnych ma charakter pozorny.
7. Rozwiązanie problemu podziału przestrzeni pozwala na otrzymanie nowego układu jednostek terytorialnych, przyporządkowanego do jednorodnego zbioru układów.
8. Uwzględnienie składników struktury wewnętrznej ekonomicznych procesów przestrzennych jest niezbędne do właściwego rozważenia problemu skali.
9. Prawidłowa identyfikacja problemu skali w warunkach autokorelacji przestrzennej wymaga ustalenia właściwej struktury przestrzennej obszaru i wykonania analiz w jej ramach.

Monografia składa się z sześciu rozdziałów.

Zarówno w przypadku problemu sposobu podziału przestrzeni, jak i problemu skali, kluczową kwestią jest proces agregacji danych. W związku z tym rozdział pierwszy pracy rozpoczyna się od teoretycznych rozważań na temat procesu agregacji. Szczegółowo przedstawiony został związek między zagadnieniem agregacji danych oraz problemem zmiennej jednostki odniesienia. Omówiono także rolę agregacji danych w badaniach ekonomicznych. Teoretyczne rozważania dotyczyły definicji agregacji danych, typów agregacji w podziale na agregację podmiotową, agregację dóbr i agregację w czasie (zob. Pawłowski, 1969, s. 234–252), budowy systemu agregacji danych oraz specyfiki agregacji danych przestrzennych.

W rozdziale drugim rozważone zostało zagadnienie poprawności przestrzennych badań ekonomicznych oraz poprawności procesu agregacji. W rozdziale przedstawiony został wyraźny podział przestrzennych badań eko-

nomicznych na trzy etapy: etap postawienia problemu badawczego, etap wykonania przestrzennej analizy ekonomicznej oraz etap rozwiązania problemu badawczego na podstawie wniosków uzyskanych z analizy.

W ramach przeprowadzonych rozważań autor wprowadził dwa ważne pojęcia: pojęcie jednorodnego przyczynowo układu danych przestrzennych oraz pojęcie jednorodnego zbioru układów. Przestrzenne badania ekonomiczne mogą być prowadzone na różnych poziomach agregacji, co wiąże się z koniecznością zbadania jednorodności przyczynowej danych przestrzennych dla każdego z tych poziomów. Rozpatrzenie tej kwestii pozwoliło autorowi na opracowanie pojęcia jednorodnego zbioru układów, który określony został w pracy jako zbiór układów jednostek terytorialnych na różnych poziomach agregacji, gdzie odniesione do tych układów dane przestrzenne posiadają własność jednorodności przyczynowej. Oznacza to, że ustalenie jednorodnego zbioru układów w ramach przestrzennych badań ekonomicznych jest koniecznym warunkiem uzyskania poprawnych wyników badań.

W ostatniej części rozdziału drugiego przedyskutowano podjęty przez Pawłowskiego problem ustalenia relacji między makroparametrem a mikroparametrami (zob. Pawłowski, 1969, 239–244). W wyniku przeprowadzonych rozważań wyprowadzone zostały równania określające relację między mikroparametrami dla trzech makroregionów a makroparametrem dla całego obszaru Polski. Zgodnie z wyznaczoną relacją ocena makroparametru stanowi kombinację liniową ocen mikroparametrów oraz odpowiednio określonych wag. Otrzymane wyniki pokrywają się z wnioskami Pawłowskiego dotyczącymi analizy szeregów czasowych.

W rozdziale trzecim przeprowadzone zostały studia literaturowe nad zagadnieniem zmiennej jednostki odniesienia, a następnie skonfrontowane z rozważaniami z rozdziału pierwszego i drugiego. Na przykładzie konkretnych prac literaturowych wskazano, na czym polegają słabości przedstawianych w nich badań nad zagadnieniem zmiennej jednostki odniesienia. W omawianych pracach dotyczących zarówno problemu sposobu podziału przestrzeni, jak i problemu skali, badania wykonywane były na podstawie wybieranych w dowolny sposób układów jednostek terytorialnych. Ma to duże znaczenie dla przeprowadzanych analiz przestrzennych, ponieważ przyjęcie pojedynczych układów danych przestrzennych, niewchodzących w skład jednorodnego zbioru układów, prowadzić może do niepoprawnych wniosków. Krytyczna ocena prac z literatury przedmiotu skonfrontowana z rozważaniami autora na temat poprawności przestrzennych badań ekonomicznych wskazała na potrzebę redefinicji tych problemów. Autor przedstawia propozycję redefinicji obydwu problemów na końcu rozdziału trzeciego.

W związku z zaproponowaną redefinicją problemu sposobu podziału przestrzeni, w rozdziale czwartym rozważono go na przykładzie przestrzennych badań ekonomicznych dotyczących rozwoju rolnictwa w Polsce na poziomie makroregionów. W wyniku wykonanych badań wyznaczony został nowy układ

czterech makroregionów rolniczych, co stanowiło rozwiązanie podjętego przez autora problemu sposobu podziału przestrzeni.

Następnie w rozdziale czwartym podjęta została ponownie kwestia ustalenia relacji między mikroparametrami a makroparametrem na przykładzie makroregionów rolniczych. W wyniku przeprowadzonych badań ustalono relację między mikroparametrami dla makroregionów rolniczych SGM a makroparametrem dla obszaru Polski, co poszerzyło postawiony przez Pawłowskiego problem o aspekty przestrzenne.

W rozdziałach piątym i szóstym dokonano analizy problemu skali na podstawie danych generowanych, której celem było sprawdzenie wpływu zmiany poziomu agregacji na otrzymywane wyniki w ramach wykonanych analiz symulacyjnych. W rozdziale piątym analizy symulacyjne przeprowadzone zostały na podstawie generowanych danych przestrzennych wyrażonych w wartościach średnich, a w rozdziale szóstym podstawę obliczeń stanowiły generowane dane przestrzenne ważone wielkością obszaru. Zarówno w rozdziale piątym, jak i w rozdziale szóstym w kolejnych scenariuszach procedury symulacyjnej zakładano wybrane własności procesów przestrzennych.

W scenariuszu pierwszym generowano realizacje procesów szumów przestrzennych o stałym poziomie wartości oczekiwanej i wariancji oraz braku autokorelacji przestrzennej. W scenariuszu drugim analizowano sytuację, gdzie w wewnętrznej strukturze procesów przestrzennych zakładano występowanie własności systematycznej niejednorodności. W ostatnim, trzecim scenariuszu założono możliwość występowania autokorelacji przestrzennej w analizowanych procesach.

Na podstawie wykonanych analiz w ramach scenariusza pierwszego i drugiego nie stwierdzono istotnych różnic w ocenach własności procesów przestrzennych oraz łączących je zależnościach przyczynowych, które byłyby wynikiem procesu agregacji. Natomiast w przypadku własności autokorelacji przestrzennej w scenariuszu trzecim ustalono występowanie problemu skali, gdzie analizowane procesy charakteryzowały się dodatnią autokorelacją przestrzenną. Oznacza to, że w sytuacji identyfikacji autozależności przestrzennych zachodzą znaczne zmiany w otrzymywanych rezultatach badań na skutek procesu agregacji danych.

Monografia zakończona została podsumowaniem, w którym przedstawione zostały wnioski z przeprowadzonych rozważań teoretycznych, jak również z wykonanych analiz symulacyjnych oraz badań empirycznych odnoszących się do przestrzennych zjawisk ekonomicznych.

Wszystkie obliczenia numeryczne zostały przeprowadzone w ramach oprogramowania R, przy wykorzystaniu gotowych pakietów lub napisanych przez autora procedur. Część testów statystycznych zostało wykonanych za pomocą narzędzi programu gretl. Natomiast rysunki przygotowane zostały w programie QGIS. Dane statystyczne pozyskane zostały z bazy Bank Danych Lokalnych udostępnianej przez Główny Urząd Statystyczny.

Praca powstała w wyniku realizacji projektu badawczego nr rej. 2015/17/B/HS4/01004 finansowanego ze środków Narodowego Centrum Nauki.

Pragnę wyrazić podziękowania Recenzentom: prof. zw. dr. hab. Janowi Zawadzkiemu oraz dr hab. Elżbiecie Szulc, prof. UMK, za wnikliwe recenzje wydawnicze, w szczególności za uwagi krytyczne, a także cenne wskazówki, których uwzględnienie przyczyniło się do podwyższenia wartości merytorycznej pracy. Pragnę wyrazić również serdeczne podziękowania Koleżankom i Kolegom z Katedry Ekonometrii i Statystyki oraz Katedry Zastosowań Informatyki i Matematyki w Ekonomii, Wydziału Nauk Ekonomicznych i Zarządzania Uniwersytetu Mikołaja Kopernika w Toruniu za wymianę doświadczeń oraz naukowe wsparcie.

