Tomasz Kossowski

Uniwersytet im. Adama Mickiewicza
Instytut Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej

Teoretyczne aspekty modelowania przestrzennego w badaniach regionalnych

Zarys treści: Modelowanie przestrzenne jest obecnie jednym z podstawowych narzędzi badawczych wykorzystywanych w analizie regionalnej. Modele przestrzenne są rozszerzeniem klasycznych modeli ekonometrycznych, do których włączane są tak zwane efekty przestrzenne: przestrzenna zależność i przestrzenna heterogeniczność. Artykuł prezentuje podstawy teoretyczne modelowania przestrzennego wraz z definicjami podstawowych pojęć oraz analizą ich własności. Przedstawione są również metody estymacji i diagnostyki modeli przestrzennych. W pracy wskazuje się też z jednej strony na złożoność modelowania przestrzennego, a z drugiej strony na użyteczność takiego podejścia badawczego. Zarysowane zostały także trendy rozwojowe modelowania przestrzennego.

Słowa kluczowe: modelowanie przestrzenne, przestrzenna zależność, przestrzenna heterogeniczność, macierz wag, estymacja modeli przestrzennych

1. Wprowadzenie

W 1970 r. Waldo Tobler sformułował pierwsze prawo geografii głoszące, że wszystko jest powiązane ze sobą, ale obiekty bliższe są bardziej powiązane niż odległe. Wynika z niego, że lokalizacja i odległość mogą wpływać na stopień zależności pomiędzy procesami lub zjawiskami w przestrzeni. Z drugiej strony, specyficzne własności miejsc albo lokalizacji mogą wpływać na zróżnicowanie przestrzenne badanych zjawisk lub procesów. W pierwszym przypadku mówi się o istnieniu zależności przestrzennej (spatial dependence) pomiędzy zjawiskami i procesami lub w samych zjawiskach i procesach, natomiast w drugim przypadku postuluje się przestrzenną heterogeniczność (spatial heterogeneity). Anselin (1988) zalicza przestrzenną zależność i przestrzenną heterogeniczność do efektów przestrzennych (spatial effects). Wynika stąd, że jednym z podstawowych zagadnień badanych na gruncie geografii powinna być analiza lub modelowanie efektów przestrzennych.

Wiele problemów rozważanych w geografii i Regional Science wymaga włączenia struktury przestrzennej zależności pomiędzy zjawiskami lub procesami występującymi w różnych punktach przestrzeni. Już w 1979 r. Paelinck i Klassen zwrócili uwagę na konieczność przestrzennego modelowania zmiennych ekonomicznych.

Tymczasem wykorzystywane często przez polskich geografów klasyczne modele regresyjne nie uwzględniają występowania efektów przestrzennych. Struktura tych modeli oraz używane metody estymacji zakładają niezależność obserwacji oraz błędów losowych, stabilność parametrów i samej modelowanej relacji, a zatem modele te z definicji są „aprzestrzenne". Ich użyteczność do modelowania zjawisk bądź procesów z występującym istotnym efektem przestrzennym jest obniżona, a uzyskane w ten sposób wyniki mogą zostać zafałszowane. Dzieje się tak, ponieważ występowanie przestrzennej zależności prowadzi do błędów w estymacji parametrów modelu ze względu na niespełnienie jego założeń, a tym samym do uzyskania niepoprawnych ocen parametrów.

Rozwiązaniem są modele przestrzenne intensywnie rozwijane od lat 70. XX w. Opublikowano wiele prac teoretycznych dotyczących modeli przestrzennych, w tym fundamentalne prace Cliffa i Orda (1973, 1981), Anselina (1988), Anselina, Floraxa i in. (1995), oraz wiele innych publikacji. Na tle piśmiennictwa światowego dorobek polskich autorów jest stosunkowo skromny. Wśród publikacji książkowych wyróżnić można trzy pozycje, traktujące przynajmniej w pewnej części o modelowaniu przestrzennym. Są to: „Ekonometria przestrzenna" pod redakcją A. Zeliasia (1991), „Ekonometria i statystyka przestrzenna z wykorzystaniem programu R CRAN" autorstwa K. Kopczewskiej (2006) oraz „Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych" pod redakcją B. Sucheckiego napisana ostatnio (2010) przez zespół ekonometryków z Uniwersytetu Łódzkiego.

W ostatnim czasie obserwuje się wzrost zainteresowania problematyką modelowania przestrzennego wśród polskich geografów ekonomicznych. Wymienić można tutaj prace Janca (2006, 2007) eksperymentujące z modelami regresji przestrzennej oraz lokalnej zależności przestrzennej czy pracę Maćkiewicz (2007) stosującą geograficznie ważoną regresję w modelowaniu cen nieruchomości niezabudowanych w aglomeracji poznańskiej. Problematykę budowy, estymacji i testowania modeli przestrzennych podejmuje Kossowski (2009) i praktycznie stosuje do modelowania dochodów własnych gmin w Polsce (Kossowski, Motek 2009). Wykorzystanie modeli przestrzennych w badaniach nad konwergencją gospodarczą komentują Ratajczak (2008) i Kossowski (2009a), natomiast próbę ich estymacji przeprowadziła Olejnik (2008).

2. Efekty przestrzenne

Modele przestrzenne uwzględniają występowanie efektów przestrzennych, do których zalicza się przestrzenną zależność i przestrzenną heterogeniczność. Przestrzenną zależnością nazywamy relację funkcyjną pomiędzy tym, co się dzieje w jednym punkcie przestrzeni, a tym, co się dzieje gdziekolwiek indziej (Anselin 1988). Dla każdego $i \in S$, gdzie S jest zbiorem jednostek przestrzennych, zapisuje się

$$y_i = f(y_1, y_2, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n).$$

Dla tak sformułowanej relacji podstawowym problemem jest identyfikowalność, a zatem możliwość jej estymacji parametrów i ich oceny. Przy n jednostkach obserwacji otrzymujemy $n(n-1)$ parametrów, gdyż po prawej stronie każdego z n równań występuje $(n-1)$ parametrów związanych z każdą z $(n-1)$ obserwacji. W jeszcze bardziej złożonym podejściu można przyjąć, że postać funkcji f jest również zależna od położenia i, co spowoduje wprowadzenie do powyższej zależności efektu przestrzennej heterogeniczności relacji. Rozwiązaniem problemu identyfikacji jest przyjęcie dodatkowych założeń dotyczących 1) postaci funkcji f, 2) zmniejszenia liczby parametrów poprzez narzucenie warunku ich stabilności przestrzennej (homogeniczności). Przyjęcie tych założeń pozwala na estymację, weryfikację i diagnostykę modelu zależności przestrzennej.

Na ogół przyjmuje się, że zależność przestrzenna może być spowodowana jedną z dwóch przyczyn, np. Anselin (1988). Po pierwsze, zależność przestrzenna może być związana z występowaniem błędów pomiarowych lub ogólniej błędów losowych w jednostkach przestrzennych. Te błędy są skutkiem niezgodności pomiędzy zasięgiem analizowanego zjawiska a podziałem obszaru na ciągłe jednostki przestrzenne. Jeżeli zjawisko występuje tylko w pewnej części jednostki przestrzennej, jego poziom oraz błąd pomiarowy agregowany jest do całej jednostki przestrzennej. Jeżeli zjawisko to występuje również w pewnej części jednostki sąsiadującej, to mówi się o pozornej zależności przestrzennej (nuisance spatial dependence). Zależność przestrzenna może również wyniknąć z agregacji danych do większych jednostek przestrzennych, a niezgodność granic zjawiska i jednostek oraz agregacja danych na ogół powoduje wzrost błędów. Problem agregacji danych w modelowaniu przestrzennym został przedyskutowany w pracy Paelincka (2000). Jeżeli błędy te mają tendencję do rozlewania się poprzez granice jednostek przestrzennych, to mówimy o efekcie spill-over.

Po drugie, zależność przestrzenna może wynikać z przestrzennego charakteru działalności człowieka. Ten typ nazywamy substancjalną zależnością przestrzenną (substantive spatial dependence). Ma on charakter fundamentalny z wyraźnym odniesieniem do przestrzennego wymiaru działalności ludzkiej.

Przestrzenną heterogenicznością (spatial heterogeneity) nazywamy niestabilność przestrzenną relacji w sensie jej postaci funkcyjnej f_i lub niestabilność przestrzenną parametrów β_{ik} badanej relacji. Przestrzenną heterogeniczność zapisuje się następująco:

$$y_i = f_i(x_{ik}, \beta_{ik}, \varepsilon_i),$$

gdzie i oznacza jednostkę przestrzenną, a ε_i błąd losowy.

Przestrzenna heterogeniczność jest skutkiem braku stacjonarności przestrzennej, tzn. nie spełnione są warunki:

$$\begin{gathered} EX(i) = EX(i+p) = \mu, \\ E(X(i))^2 = E(X(i+p))^2 = \sigma^2, \\ E(X(i)X(j)) = \gamma(d(i,j)). \end{gathered}$$

Występowanie przestrzennej heterogeniczności może wynikać z zaburzeń stabilności przestrzennej średniej, wariancji lub autokowariancji. W przypadku braku przestrzennej heterogeniczności funkcja autokowariancji zależy tylko od odległości. Metody badania i modelowania przestrzennej heterogeniczności nie będą jednak w tej pracy omawiane.

3. Macierze wag przestrzennych

Macierz wag przestrzennych jest formalnym wyrażeniem zależności przestrzennej pomiędzy jednostkami obserwacji (Anselin 1988). Określa ona strukturę przestrzenną sąsiedztwa oraz mierzy siłę potencjalnych interakcji. Do tej pory nie ustalono jednolitego poglądu, jak powinno się zapisywać strukturę przestrzenną sąsiedztwa. Stąd określono wiele sposobów kodowania elementów macierzy wag przestrzennych. W ogólności, sąsiedztwo takie można zdefiniować następująco: najbliższymi sąsiadami są obiekty mające wspólną granicę lub mieszczące się w otoczeniu o określonym promieniu g (Suchecki i in. 2010).

Najbardziej elementarnym typem macierzy wag przestrzennych jest zero-jedynkowa macierz bezpośredniego (najbliższego) sąsiedztwa **C**. Elementy tej macierzy $c_{ij} = 1$, jeżeli i sąsiaduje z j albo $c_{ij} = 0$, jeżeli i nie sąsiaduje z j. W tym przypadku sąsiedztwo oznacza fakt posiadania wspólnej granicy. Macierz **C** jest macierzą symetryczną, której elementy leżące na głównej przekątnej są równe zero (ponieważ nie można być swoim sąsiadem).

Przekształcając elementy macierzy bezpośredniego sąsiedztwa **C** według formuły:

$$w_{ij} = \frac{c_{ij}}{\sum_{j=1}^{n} c_{ij}},$$

otrzymuje się macierz **W**, której suma elementów w każdym wierszu wynosi 1, lecz traci ona symetrię. Standaryzowana macierz **W** jest najczęściej wykorzystywaną w analizach macierzą wag przestrzennych. Jak zauważają Tiefelsdorf i in. (1998), w wierszach standaryzowanej macierzy **W** zbyt duże wagi są przypisywane do jednostek z małą liczbą sąsiadów, co prowadzi do przeszacowania stopnia interakcji przestrzennej. Problem ten dotyczy z reguły jednostek leżących na granicach badanego obszaru (efekt granicy).

Wymienione powyżej schematy kodowania należą do najpopularniejszych. Getis i Aldstadt (2004) podają 11 sposobów konstrukcji macierzy wag. Griffith (1996) przeprowadził krytyczny przegląd macierzy wag przestrzennych. Autor ten wprowadził 5 zaleceń dotyczących wyboru właściwej macierzy: 1) uwzględnienie macierzy wag powinno prowadzić do otrzymania lepszych wyników niż zakładanie przestrzennej niezależności, 2) podział przestrzeni na jednostki powinien być pomiędzy regularną siatką kwadratów a siatką heksagonów, dla przypadku podziałów nieregularnych liczba sąsiadów powinna wynosić od 4 do 6, 3) minimalna liczba jednostek przestrzennych to 60, 4) modele przestrzenne z niższym rzędem sąsiedztw powinny być preferowane przed modelami z wyższymi rzędami, 5) lepsze są macierze o mniejszej liczbie sąsiadów niż za dużej liczbie sąsiadów.

Opisane tutaj własności dotyczą egzogenicznych macierzy wag, które są ustalane poza modelem, na podstawie przyjętych założeń. W ostatnich latach nabiera znaczenia problem endogenizacji macierzy wag, a co za tym idzie – poszukiwanie sposobów jej estymacji. Przykładowe próby endogenizacji macierzy wag podjęto w pracach Ouda i in. (2010) oraz Vazqueza (2010).

4. Badanie efektów przestrzennych

Badanie efektów ma na celu ustalenie typu efektu przestrzennego (przestrzennej zależności lub przestrzennej heterogeniczności) występującego w danych przestrzennie zlokalizowanych. Zależność przestrzenna bardzo często jest utożsamiana z autokorelacją przestrzenną (Anselin 1988), przez co testowanie występowania zależności przestrzennej sprowadza się do weryfikacji hipotezy o istnieniu autokorelacji przestrzennej w danych przestrzennie zlokalizowanych.

Autokorelacja przestrzenna jest to korelacja pomiędzy wartościami jednej zmiennej pomierzonymi w różnych punktach przestrzeni. Definicja autokorelacji przestrzennej wprowadza odchylenie od założenia niezależności obserwacji, jakie często jest formułowane w klasycznej statystyce. Odpowiednikiem autokorelacji przestrzennej jest autokorelacja w czasie, gdzie wartość zmiennej w danej chwili t może zależeć od obserwacji wcześniejszych. Autokorelacja przestrzenna ma charakter wielokierunkowy, podczas gdy autokorelacja czasowa jest jednokierunkowa.

Miary autokorelacji przestrzennej są wykorzystywane od połowy XX w. Najlepiej poznaną miarą autokorelacji przestrzennej jest współczynnik I Morana (1950). Jest on zdefiniowany następująco:

$$I = \frac{n}{s_0} \cdot \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2},$$

gdzie n jest liczbą jednostek przestrzennych, c_{ij} oznacza element zero-jedynkowej macierzy wag $\mathbf{C}$, s_0 jest sumą elementów macierzy $\mathbf{C}$, x_i oznacza wartość obserwacji w jednostce i.

Przyjmuje się, że jeżeli wartość współczynnika I Morana jest większa od $-\frac{1}{n-1}$, to mamy dodatnią autokorelację przestrzenną, a dla I mniejszego od tej wartości mamy ujemną autokorelację przestrzenną. Dla współczynnika wynoszącego w przybliżeniu $-\frac{1}{n-1}$ przyjmuje się, że rozkład wartości zmiennej x w przestrzeni jest losowy.

Na rycinie 1 przedstawiono przykładowe układy przestrzenne i odpowiadające im wartości współczynnika autokorelacji przestrzennej dla pewnej sztucznej zmiennej X. Dodatnia autokorelacja przestrzenna oznacza tendencję do klastrowania się jednostek przestrzennych, natomiast ujemna autokorelacja przestrzenna powoduje powstawanie układów mozaikowych. Autokorelacja przestrzenna bliska zeru generuje układy losowe.

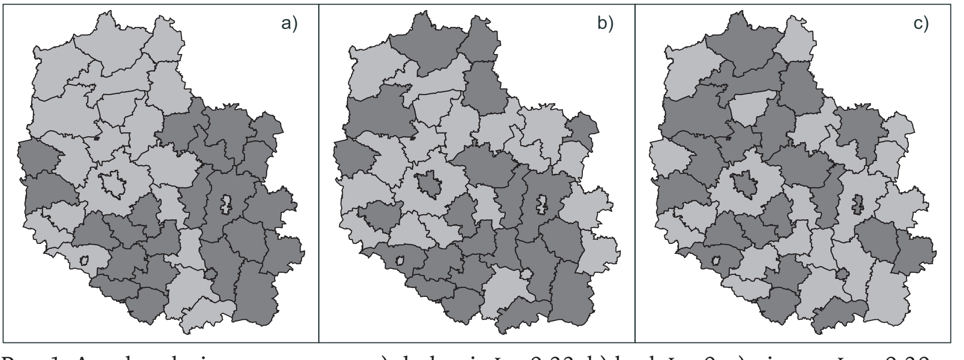
W praktyce współczynnik Morana zapisuje się w krótkiej postaci macierzowej:

$$I = \frac{\mathbf{z}'\mathbf{W}\mathbf{z}}{\mathbf{z}'\mathbf{z}},$$

w której $\mathbf{W} = \frac{n}{s_0}\mathbf{C}$, $z_i = x_i - \bar{x}$, a zatem $\mathbf{W}$ jest wierszowo standaryzowaną macierzą wag przestrzennych. Współczynnik autokorelacji przestrzennej Morana przyjmuje zasadniczo wartości z przedziału [–1,1], ale z teoretycznego punktu widzenia ten przedział może być inny. Dzieje się tak, ponieważ w przeciwieństwie do współczynnika korelacji Pearsona czy Spearmana, indeks Morana jest nieunormowany. Przyjmijmy, że $\mathbf{M} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'$, gdzie $\mathbf{1}$ jest wektorem jedynek długości n. Jak zostało ustalone w pracy de Jonga i in. (1984), wartości graniczne współczynnika autokorelacji przestrzennej Morana są następujące:

$$\lambda_{\min}(\mathbf{MWM}) \leq I \leq \lambda_{\max}(\mathbf{MWM}),$$

gdzie $\lambda_{\min}$, $\lambda_{\max}$ są najmniejszymi i największymi wartościami własnymi macierzy **MWM**. Wektory własne odpowiadające $\lambda_{\min}$, $\lambda_{\max}$ generują maksymalną i minimalną wartość współczynnika Morana dla danej macierzy wag (Griffith 2003). Odpowiadające tym wartościom własnym wektory własne macierzy **MWM** generują maksymalne i minimalne wartości współczynnika Morana. Problemem jest wyznaczanie wartości własnych dla dużych macierzy **MWM**. W ostatnim czasie Griffith (2000, 2004) wyprowadził wzory funkcji własnych (eigenfuctions) pozwalających w szybki sposób wyliczać wartości własne macierzy **MWM** dla regularnych po-

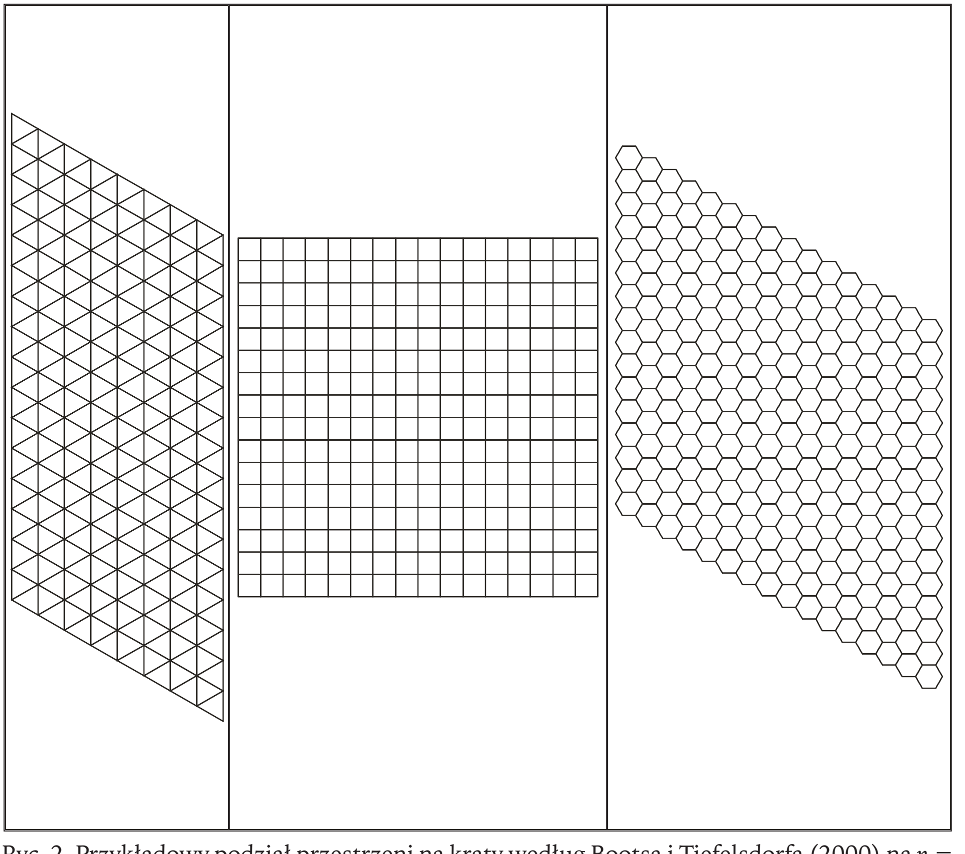


Ryc. 1. Autokorelacja przestrzenna: a) dodatnia I = 0,33, b) brak I = 0, c) ujemna I = –0,30
Źródło: Kossowski (2009).

działów przestrzeni (tzn. na trójkąty, kwadraty i sześciokąty). Dla takich podziałów przestrzeni Boots i Tiefelsdorf (2000) badali graniczne wartości współczynnika Morana.

Wyliczone przez Bootsa i Tiefelsdorfa wartości graniczne współczynnika Morana dla przedstawionych na rycinie 2 typów krat są zamieszczone w tabeli 1. Graniczne wartości współczynnika autokorelacji Morana różnią się dla analizowanych krat, ponieważ zależą od macierzy wag przestrzennych **W**. Dla każdego podziału przestrzeni wartości graniczne mogą być inne, a to oznacza, że nie ma możliwości porównywania wartości współczynnika Morana dla różnych struktur przestrzennych.

Analiza tabeli 1 pozwala zauważyć, że podział przestrzeni na heksagony (sześciokąty) powoduje, że przedział wartości współczynnika Morana jest najwęższy. Podział na heksagony ogranicza wystąpienie ujemnej autokorelacji przestrzennej. Potwierdzony jest w ten sposób wymieniony wcześniej postulat Griffitha, że podział przestrzeni powinien być bliski podziałowi na sześciokąty.



Ryc. 2. Przykładowy podział przestrzeni na kraty według Bootsa i Tiefelsdorfa (2000) na $n = 256$ elementów

Tabela 1. Wartości graniczne współczynnika Morana dla wybranych podziałów na kraty

Typ kraty	Liczba elementów kraty	Przedział wartości współczynnika Morana
Trójkątna	64	[–1.0725, 1.0330]
	256	[–1.0477, 1.0375]
	1024	[–1.0273, 1.0246]
Kwadratowa	64	[–1.0739, 0.9747]
	256	[–1.0485, 1.0216]
	1024	[–1.0276, 1.0206]
Heksagonalna	64	[–0.5519, 1.0065]
	256	[–0.5330, 1.0403]
	1024	[–0.5186, 1.0306]

Źródło: Boots, Tiefelsdorf (2000).

Hauke i Kossowski (2008) badali uogólniony podział przestrzeni za pomocą trójkątów, kwadratów i heksagonów. Jednostką bazową takiego podziału przestrzeni jest heksagon, którego boki obudowane zostały kwadratami oraz trójkątami równobocznymi (por. ryc. 3a). Otrzymana w ten sposób jednostka bazowa składa się z 13 elementów: 1 heksagonu, 6 kwadratów i 6 trójkątów równobocznych. Obrys jednostki bazowej ma kształt dwunastokąta foremnego. Na rycinie 3a oraz 3b pokazane jest składanie jednostek bazowych w bardziej złożone struktury. Łączenie elementów bazowych odbywa się poprzez łączenie ich krawędzi, a wolne miejsca wypełniane są przez trójkąty równoboczne.

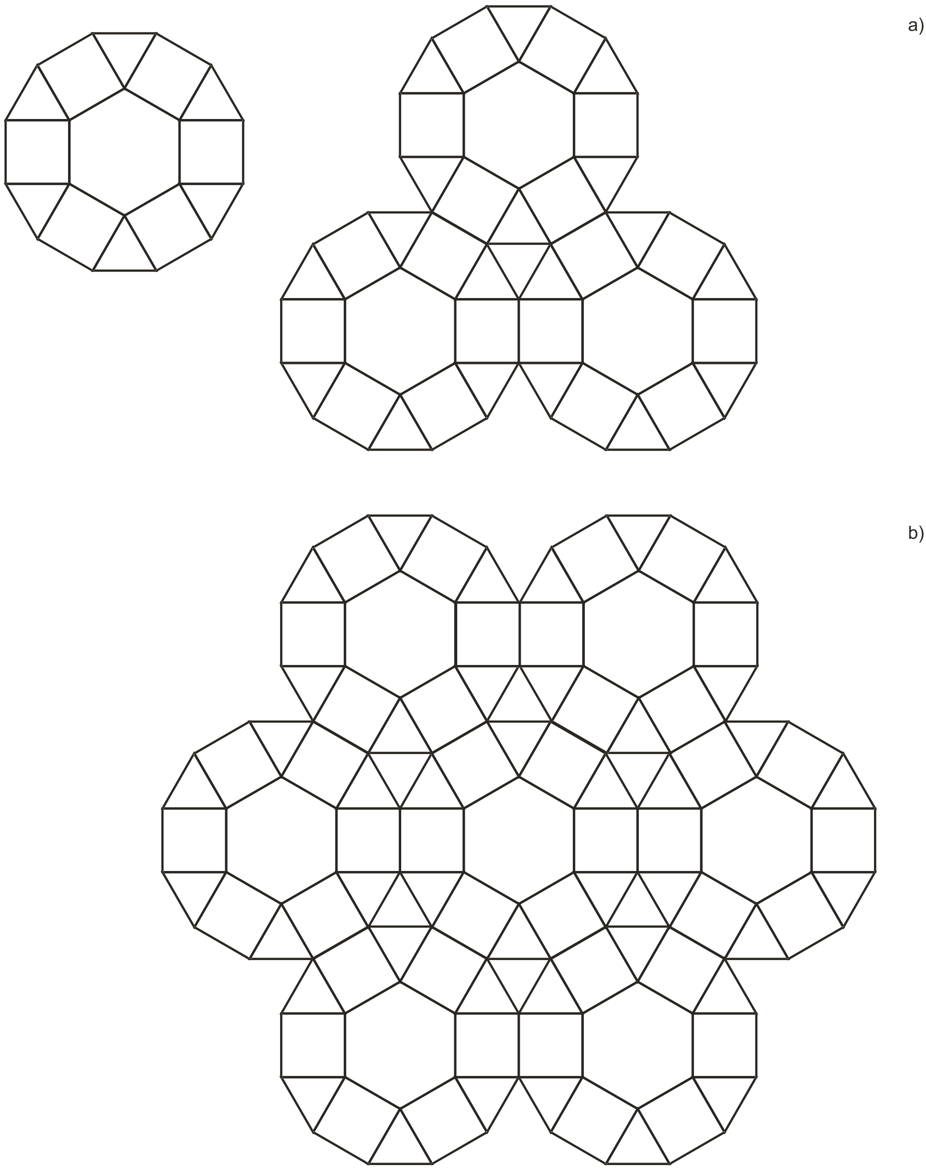
Hauke i Kossowski (2008) wyliczyli wartości graniczne współczynnika Morana dla układu zbudowanego z trzech i siedmiu elementów bazowych. W przypadku trzech elementów bazowych struktura liczy 40 regularnych figur, a przedział współczynnika wynosi [–1.1042, 1.0451], a dla siedmiu elementów bazowych jest 97 figur geometrycznych. Przedział współczynnika Morana to [–1.0978, 1.0901].

Współczynnik autokorelacji przestrzennej I Morana może zostać wykorzystany do weryfikacji hipotezy głoszącej o braku statystycznej istotności autokorelacji przestrzennej. Statystyka testowa jest postaci:

$$Z(I) = \frac{I - \mathrm{E}(I)}{\sqrt{\mathrm{Var}(I)}},$$

gdzie $\mathrm{E}(I) = -\frac{1}{n-1}$ jest wartością oczekiwaną statystyki Morana, a $\mathrm{Var}(I)$ jest jej wariancją. Postać wariancji jest złożona, jej sformułowanie znajduje się np. w pracy pod redakcją Sucheckiego (2010) i zależy od przyjętego założenia odnośnie do losowości próby. Możliwe są bowiem trzy założenia: 1) normalność – elementy próby są realizacjami zmiennych o identycznym rozkładzie normalnym, 2) losowość – wartość statystyki jest jedną z możliwych $n!$ wartości odpowiadających losowemu rozmieszczeniu wartości zmiennej w jednostkach przestrzennych, z pominięciem rozkładu zmiennych w lokalizacjach, 3) empirycznej oceny częstości rozkładu na podstawie analizy permutacji.

W przypadku 1) i 2) przyjmuje się, że statystyka *Z(I)* ma rozkład asymptotycznie normalny (Cliff i Ord 1973), ale jej zbieżność do tego rozkładu jest powolna. Jeszcze w latach 90. Tiefelsdorf i Boots (1995) wyprowadzili dokładny rozkład statystyki Morana dla małych prób. Z powodu powolnej zbieżności statystyki do rozkładu normalnego najczęściej wykorzystuje się założenie 3) prowadzące do per-



Ryc. 3. Uogólniony podział przestrzeni według Hauke i Kossowskiego (2008), a) element bazowy i prosta struktura z trzech elementów bazowych, b) struktura zbudowana z siedmiu elementów bazowych

mutacyjnej wersji tego testu. Polega ona na wykonaniu ustalonej liczby (z reguły większej od n) permutacji wartości zmiennej X na zbiorze jednostek przestrzennych. Następnie dla każdej permutacji wyliczane są współczynniki autokorelacji przestrzennej, na podstawie których buduje się rozkład empiryczny statystyki Morana. Niestety, wadą takiego podejścia jest to, że przy każdej serii permutacji otrzymuje się nieco inny rozkład empiryczny. Dla pewnych szczególnych wartości współczynnika Morana może to utrudniać ustalenie jego poziomu istotności. W przypadku testu permutacyjnego wylicza się poziom pseudoistotności p na podstawie ilorazu $p = \frac{NS+1}{NP+1}$, gdzie NS jest liczbą obliczonych wartości testu większych lub równych od rzeczywistej wartości statystyki Morana, a NP jest liczbą permutacji.

Mimo że miara autokorelacji przestrzennej zaproponowana przez Morana ma już 60 lat, ciągle znajduje się w kręgu zainteresowań badaczy. Spośród licznych prac traktujących o tym zagadnieniu, można wskazać na artykuł Lee (2001) o próbie integracji współczynnika Morana ze współczynnikiem korelacji Pearsona. Z kolei Bivand i in. (2008) analizowali moc statystyki Morana. Ostatnio Griffith (2010) napisał pracę podsumowującą wyniki dotyczące rozszerzenia własności rozkładu statystyki Morana na zmienne losowe nie mające rozkładu normalnego. Mniejszą popularnością cieszą się badania koncentrujące się na własnościach współczynnika autokorelacji przestrzennej Geary'ego (1954) oraz ten współczynnik wykorzystujące. Miara autokorelacji przestrzennej C Geary'ego jest zdefiniowana według formuły:

$$C = \frac{n-1}{2s_0} \cdot \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}(z_i - z_j)^2}{\sum_{i=1}^{n} z_i^2}.$$

Jeżeli $C < 1$, to mamy dodatnią autokorelację przestrzenną, a jeżeli $C > 1$, to występuje ujemna autokorelacja, natomiast dla $C = 1$ wartości zmiennej w przestrzeni są rozmieszczone losowo. Za pomocą zapisu macierzowego, można współczynnik Geary'ego zapisać w postaci (por. Griffith 2003):

$$C = \frac{n}{n-1}\left[\frac{n}{s_0} \cdot \frac{\mathbf{z}'\mathrm{diag}(c_{i\cdot})\mathbf{z}}{\mathbf{z}'\mathbf{z}} - I\right],$$

gdzie I jest współczynnikiem Morana. Z powyższej formuły wynika, że współczynnik Geary'ego daje się wyrazić za pomocą współczynnika Morana. Taka konstrukcja współczynnika powoduje, że C Geary'ego jest bardziej wrażliwe na występowanie lokalnych odchyleń od globalnego wzorca autokorelacji niż współczynnik Morana.

Wartości współczynnika Geary'ego zawierają się zasadniczo w przedziale [0,2], ale podobnie jak dla współczynnika Morana można wyliczyć dokładne wartości maksymalne i minimalne, korzystając z cytowanej pracy de Jonga i in. (1984). Podobnie jak dla współczynnika Morana, można wykorzystać współczynnik C do testowania istotności autokorelacji przestrzennej. Również w tym przypadku statystyka $Z(C)$ ma rozkład asymptotycznie normalny, o wartości oczekiwanej równej 1 i wariancji zależnej od założenia o losowości próby. Współczynniki autokorelacji przestrzennej Morana i Geary'ego są specjalnymi przypadkami statystyki Γ (por. Suchecki 2010). Innymi współczynnikami wykorzystywanymi do badania autokorelacji przestrzennej, choć mniej popularnymi, są statystyki joint-count (Cliff i Ord 1973, Ratajczak 1980, Kossowski 2006) oraz statystyka G Getisa-Orda (1992). W pracy z 2007 r. Arthur Getis odwołuje się do źródła pojęcia autokorelacji przestrzennej i podaje dziesięć miar dla autokorelacji.

5. Modele przestrzenne i ich estymacja

W modelowaniu przestrzennym uwzglednia się możliwe występowanie efektów przestrzennych: przestrzennej zależności lub przestrzennej heterogeniczności. Wśród modeli przestrzennych uwzględniających element przestrzennej zależności najpopularniejszy jest ogólny model przestrzennej regresji liniowej. Specyfikacja tego modelu w dość ogólnej postaci zawarta jest w pracy Anselina (1988):

$$\mathbf{y} = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$
$$\boldsymbol{\varepsilon} = \lambda \mathbf{W}_2\, \boldsymbol{\varepsilon} + \boldsymbol{\mu},$$
$$\boldsymbol{\mu} \sim N(0, \boldsymbol{\Omega}), \Omega_{ii} = h_i(\mathbf{z}', \boldsymbol{\alpha}),\ h_i > 0,$$

gdzie y jest wektorem zmiennej zależnej o wymiarach k na 1, $\boldsymbol{\beta}$ jest wektorem parametrów zmiennych egzogenicznych o wymiarze k na 1, $\mathbf{X}$ jest macierzą zmiennych egzogenicznych o wymiarach n na k, natomiast ρ, λ są współczynnikami autoregresyjnymi opóźnionej zmiennej zależnej i przestrzennie opóźnionego błędu losowego. Macierze $\mathbf{W}_1$, $\mathbf{W}_2$ są macierzami wag przestrzennych stopnia n. Elementy na głównej przekątnej macierzy kowariancji $\boldsymbol{\Omega}$ są pewną nieujemną funkcją h_i, zależną od wektora zmiennych $\mathbf{z}$ oraz parametrów $\boldsymbol{\alpha}$. Model uwzględnia możliwość wystąpienia heteroskedastycznego błędu losowego. Przy założeniu $\boldsymbol{\alpha} = 0$ otrzymujemy sytuację błędu homoskedastycznego, tzn. $h_i = \sigma^2$.

Na podstawie powyższych równań można zbudować wektor nieznanych parametrów ogólnego modelu przestrzennej regresji liniowej $\boldsymbol{\theta} = [\rho, \boldsymbol{\beta}', \lambda, \sigma^2, \boldsymbol{\alpha}']$, w którym $\boldsymbol{\beta}'$ ma k elementów, a $\boldsymbol{\alpha}'$ zawiera p elementów. Wektor $\boldsymbol{\theta}$ zawiera $3 + k + p$ nieznanych parametrów modelu. Ograniczając wartości elementów wektora parametrów, otrzymuje się zredukowane wersje modelu:

1) klasyczny model regresji liniowej LRM: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \rho = \lambda = 0,\ \boldsymbol{\alpha} = 0$, który ma $p + 2$ parametrów,
2) model opóźnienia przestrzennego SAR: $\mathbf{y} = \rho\, \mathbf{W}_1\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \lambda = 0,\ \boldsymbol{\alpha} = 0$, który ma $p + 1$ parametrów,
3) model błędu przestrzennego SEM: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} = \lambda \mathbf{W}_2 + \boldsymbol{\mu},\ \rho = 0, \boldsymbol{\alpha} = 0$, który ma również $p + 1$ parametrów,

4) mieszany regresyjno-przestrzennie autoregresyjny model z autoregresyjnym błędem losowym SAC: $\mathbf{y} = \rho\, \mathbf{W}_1 \mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, gdzie $\boldsymbol{\varepsilon} = \lambda \mathbf{W}_2\, \boldsymbol{\varepsilon} + \boldsymbol{\mu}$, ale $\boldsymbol{\mu} \sim \mathrm{N}(0,\boldsymbol{\Omega})$ $\Omega_{iii}=\sigma^2$. Model ten ma p parametrów, gdyż $\boldsymbol{\alpha} = 0$. Macierze wag $\mathbf{W}_1$, $\mathbf{W}_2$ najczęściej są równe. Model SAC jest hybrydą modeli SAR oraz SEM.

Model opisany w punkcie 1) można estymować klasyczną metodą najmniejszych kwadratów (OLS). Wyprowadzone w ten sposób estymatory parametrów są zgodne, nieobciążone oraz najbardziej efektywne. Problemem w modelowaniu przestrzennym jest to, że estymatory po włączeniu efektu przestrzennej zależności tracą swoje „dobre" własności. Dla sytuacji modelu 2), czyli wystąpienia opóźnionej przestrzennie zmiennej zależnej, estymatory nie są zgodne i obciążone (Anselin 1988). Natomiast dla modelu 3), a więc w przypadku wystąpienia autokorelacji przestrzennej błędu losowego, estymatory te są nieobciążone i nieefektywne. W związku z tym do estymacji parametrów modelu 2) oraz 3) używa się innych metod estymacji niż OLS.

Dla modeli typu SAC wyprowadzono estymatory największej wiarogodności, których szczególnymi przypadkami są estymatory dla modeli SAR i SEM. Wyprowadzenie tych estymatorów można znaleźć w pracach Cliffa i Orda (1981), Anselina (1988) oraz w polskiej pracy zbiorowej pod redakcją Sucheckiego (2010). Stosowanie tej metody niesie jednak ze sobą pewne trudności, z których jedną jest poszukiwanie wartości własnych dla dużych i rzadkich macierzy. Ograniczenie to może zostać pominięte w przypadku estymacji parametrów uogólnioną metodą momentów (GMM). Estymatory uogólnionej metody momentów zostały wyprowadzone przez Kelejiana i Pruchę (1999). Metoda ta jest jednak bardziej restrykcyjna, gdyż wymaga przyjęcia dodatkowych czterech założeń (por. Suchecki 2010). Porównanie numerycznej efektywności obu metod zawiera praca Walde i in. (2008).

Oprócz wyżej wymienionych metod estymacji, ostatnio rozwinięta została również podwójna metoda najmniejszych kwadratów (Kelejian i Prucha 1997, 1998, Lee 2007). Haining (1978) i Bivand (1984) zaproponowali wykorzystanie metody zmiennych instrumentalnych (IV) do estymacji modeli przestrzennych, ale nie jest ona w pracach tych autorów używana. Podstawy teoretyczne metody zmiennych instrumentalnych, a także podejścia bayesowskiego przedstawia Anselin (1988).

6. Specyfikacja modelu przestrzennego, jego oszacowanie i diagnostyka

Specyfikacja, oszacowanie oraz diagnostyka modelu przestrzennego jest procedurą złożoną z kilku etapów. Są one następujące:

1. Redukcja zbioru zmiennych objaśniających przy użyciu klasycznego modelu liniowego szacowanego przez metodę najmniejszych kwadratów: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$, $\varepsilon \sim \mathrm{N}(\mu, \sigma^2\mathbf{I})$, z którego eliminuje się nieistotne statystycznie zmienne.
2. Testowanie autokorelacji przestrzennej reszt modelu liniowego oszacowanego zredukowanej liczby zmiennych. Na ogół modele te szacuje się na podstawie:

a) wartości zmiennych albo b) logarytmów wartości zmiennych. Jak podają Cliff i Ord (1970), występowanie autokorelacji przestrzennej może być skutkiem efektu szacowania zależności nieliniowej modelem liniowym. W testowaniu autokorelacji przestrzennej wykorzystywana jest statystyka *I* Morana w postaci (Cliff i Ord 1972, 1973, Anselin i Florax 1995) $I = \frac{\mathbf{e}'\mathbf{W}\mathbf{e}}{\mathbf{e}'\mathbf{e}}$, gdzie **e** jest wektorem reszt z regresji. Statystyka ma rozkład asymptotycznie normalny. Jeżeli wartość sprawdzianu Morana jest istotna statystycznie, to stosuje się podejście przestrzenne.

3. Badanie typu zależności przestrzennej pozwalające na wybór odpowiedniej specyfikacji modelu przestrzennego, uwzględniającego: a) zależność przestrzenną zmiennej objaśnianej, b) zależność przestrzenną błędu losowego, c) zależność przestrzenną zmiennej zależnej i błędu losowego. Sprawdzanie typu zależności przeprowadza się w oparciu o testy *LM* (Lagrange Multiplier). Mają one zastosowanie do klasycznego modelu regresji oszacowanego metodą najmniejszych kwadratów. Test LM_ρ (Anselin 1988, 1995) sprawdza, czy model powinien uwzględniać sytuację a), natomiast test LM_λ (Burridge 1980) jest wykorzystywany do sprawdzenia przypadku b). Wersje odporne tych testów, takie jak RLM_ρ i RLM_λ (Bera i Yoon 1992), są niewrażliwe na lokalne błędy w specyfikacji modelu. Omawiane statystyki mają rozkład χ^2 z jednym stopniem swobody. Zazwyczaj przeprowadza się serię czterech powyższych testów, przy czym zgodnie z tzw. podejściem klasycznym, rozstrzygający jest test najbardziej istotny (Kossowski, Motek 2009). Szczegółowy opis własności tych testów podają Anselin i Florax (1995), Anselin i in. (1996) oraz Bera i Yoon (1992). Anselin i Moreno (2003) porównują własności testów dla składników błędu przestrzennego.
   W przypadku sytuacji c) używa się statystyki *SARMA* (Anselin 1988a, 1994) będącej sumą statystyki odpornej *LM* dla opóźnienia przestrzennego i *LM* dla błędu przestrzennego. Jej rozkład to χ^2 z dwoma stopniami swobody. Istotne wartości statystyki *SARMA* pozwalają na przyjęcie modelu z zależnością przestrzenną zmiennej y i średnią ruchomą błędu losowego[1].
4. Estymacja parametrów modelu odbywa się z użyciem metod opisanych we wcześniejszych punktach pracy.
5. Diagnostyka modelu polega na: 1) testowaniu założeń modelu, 2) testowaniu parametrów strukturalnych oraz całego modelu, 3) badaniu siły informacyjnej modelu oraz jego dopasowaniu. W sprawdzaniu założeń modelu stosuje się klasyczne testy ekonometryczne, takie jak: test normalności rozkładu reszt Jarque'a-Bery, heteroskedastyczności Breuscha-Pagana w wersji dla modeli zwykłych i przestrzennych, test White'a dla heteroskedastyczności modelu oraz test ilorazu wiarogodności *LR* dla heteroskedastyczności grupowej[2]. Przy testowaniu parametrów strukturalnych modelu zastosowanie mają klasyczne testy *t* oraz test Walda dla parametru autoregresyjnego. Badanie siły informa-

[1] Model SARMA zawiera składnik zależności przestrzennej zmiennej objaśnianej oraz przestrzenną średnią ruchomą błędu losowego, natomiast w modelu SARAR (oznaczany również SARSAR) błąd losowy jest autoregresyjny.

[2] Test ten potwierdza zasadność uwzględnienia efektów przestrzennych w modelu.

cyjnej modelu opiera się na wykorzystaniu kryteriów informacyjnych, do których należą AIC (kryterium informacyjne Akaike'a), BIC (kryterium bayesowskie), kryterium Schwarza oraz logLik (logarytm ilorazu wiarogodności). Jeżeli porównuje się dwa modele, to model lepszy ma najniższe AIC i BIC oraz najwyższe logLik. Jak podają Anselin (1988) i Maddala (2006), testy Walda (*W*), *LR* oraz *LM* są asymptotycznie równoważne i przy poprawnej specyfikacji modelu zachodzi nierówność $W \geq LR \geq LM$. Testy Walda, *LR*, *LM* oraz kryterium AIC dla modeli przestrzennych zostały zamieszczone w pracy Anselina (1988).

Procedura budowy i testowania modelu przestrzennego jest bardziej złożona niż w przypadku klasycznym. Przedstawione powyżej reguły postępowania odnoszą się do przypadków najprostszych, tzn. szacowania i weryfikacji modeli typu SAR lub SEM. Pojawienie się heteroskedastyczności albo heterogeniczności przestrzennej w przypadku jednoczesnej obecności autokorelacji przestrzennej wymusza stosowanie metod i testów jeszcze bardziej skomplikowanych. Ich krótki przegląd można znaleźć w pracy pod redakcją Sucheckiego (2010).

7. Zakończenie

W niniejszym opracowaniu zaprezentowano zagadnienie analizy efektów przestrzennych, których występowanie jest przyczynkiem do stosowania zmodyfikowanych modeli ekonometrycznych w analizie regionalnej. Wprowadzenie oddziaływania tych efektów z reguły poprawia wartość poznawczą oraz moc objaśniającą zastosowanych modeli, czego liczne dowody można znaleźć w literaturze światowej. Pamiętać jednak należy, że modelowanie przestrzenne nie jest zagadnieniem prostym. Podstawowym problemem są niedoskonałe metody rozróżniania efektów przestrzennych pomiędzy sobą, to jest przestrzennej zależności i przestrzennej heterogeniczności. Ponadto metody estymacji takich modeli są bardziej złożone niż modeli klasycznych. Niektóre z nich wymagają specjalnych założeń dotyczących modelu (jak metoda GMM) bądź są wrażliwe na błędy w specyfikacji modelu (np. występowanie heteroskedastycznego błędu losowego). Inne z kolei wymagają znacznych mocy obliczeniowych, np. metoda największej wiarogodności, która wykorzystuje wyznaczniki macierzy rzadkich. Obecnie badania zmierzają w kierunku opracowania jak najlepszych estymatorów dla parametrów modeli przestrzennych, także tych bardziej skomplikowanych. Prowadzi się również badania nad estymacją modeli czasowo-przestrzennych. Osobny nurt badawczy, stosunkowo niedawno zapoczątkowany, wiąże się z problematyką endogenizacji, a co za tym idzie – estymacji macierzy wag. Problem ten jest związany z odrzuceniem założenia o egzogeniczności macierzy wag względem modelu.

Z drugiej strony, nie ma przeszkód, aby stosować modele przestrzenne w analizie regionalnej. Od kilkunastu lat trwa dynamiczny rozwój oprogramowania komputerowego pozwalającego skutecznie wdrażać tego rodzaju analizy. Wystarczy wspomnieć o projekcie R, w ramach którego rozwijany jest pakiet *spdep* przeznaczony właśnie do modelowania przestrzennego. Do oprogramowania komercyjnego dopisywane są odpowiednie rozszerzenia umożliwiające prowadzenie modelowania przestrzennego, np. pakiet metod ekonometrii przestrzennej napisany przez J. LeSage dla programu Matlab czy też dodatkowe procedury w programie ArcGIS. Istnieją również programy przeznaczone do budowy prostych modeli przestrzennych, których doskonałym przykładem jest Geoda opracowana przez L. Anselina. W bliskiej przyszłości dostępny będzie program PySAL napisany przez L. Anselina i S. Reya, będący rozwinięciem Geody.

Modelowanie przestrzenne stanowi jeden z głównych nurtów rozwijanych i stosowanych w analizie regionalnej metod i technik badawczych. Szczegółowy przegląd wszystkich prac z tego zakresu jest niemożliwy ze względu na ich bardzo dużą liczbę. Przykładowo modelowanie przestrzenne jest wykorzystywane w analizie produktywności regionalnej i konwergencji dochodów w Niemczech (Kosfeld i in. 2006), do badania gospodarek „wschodzących" (Kelejian i in. 2006) bądź też regionalnego wzrostu gospodarczego gospodarek dynamicznych (Chiny wg Ying 2003). Sporo jest prac wykorzystujących modele przestrzenne w zagadnieniach geografii zdrowia, np. dostępności opieki ambulatoryjnej w USA (Mobley i in. 2006) lub śmiertelności powodowanej przez nowotwory (Haining 1995). Mimo wielu zastosowań modelowanie przestrzenne nie jest wciąż popularne w praktyce planistycznej. Biorąc pod uwagę jego przydatność potwierdzoną w wielu pracach, a których dokonany tutaj przegląd jest cząstkowy, należy sądzić, że upowszechni się również w dziedzinie planowania.

Literatura

Anselin L. 1988. Spatial Econometrics: Methods and Models. Kluwer, Dordrecht.

Anselin L. 1988a. Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity. Geographical Analysis, 20: 1–17.

Anselin L. 1994. Testing for Spatial Dependence in Linear Regression Models: A Review. West Virginia University, Morgantown, Regional Research Institute Research Paper, s. 94–116.

Anselin L., Florax R.J.G.M. (red.) 1995. New Directions in Spatial Econometrics. Springer-Verlag, Berlin.

Anselin L., Bera A.K., Florax R.J.G.M., Yoon M.J. 1996. Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics, 26: 77–104.

Anselin L., Moreno R. 2003. Properties of tests for spatial error components. Regional Science and Urban Economics, 33: 595–618.

Bera A., Yoon M. 1992. Simple Diagnostic Tests for Spatial Dependence. University of Illinois, Urbana-Champaign.

Bivand R.S. 1984. Regression Modelling with Spatial Dependence: An Application of Some Class Selection and Estimation Methods. Geographical Analysis, 16: 25–37.

Bivand R. 2008. Power calculations for global and local Moran's I. Computational Statistics and Data Analysis, 53, 8: 2859–2872.

Boots B., Tiefelsdorf M. 2000. Global and local spatial autocorrelation in bounded regular tessalations. Journal of Geographical Systems, 2: 319–348.

Burridge P. 1980. On the Cliff-Ord Test for Spatial Autocorrelation. Journal of The Royal Statistical Society B, 42: 107–108.

Cliff A.D., Ord J.K. 1970. Spatial Autocorrelation: A Review of Existing and New Measures with Applications. Economic Geography, 46: 269–272.
Cliff A.D., Ord J.K. 1972. Testing for Spatial Autocorrelation Among Regression Residuals. Geographical Analysis, 4: 267–284.
Cliff A.D., Ord J.K. 1973. Spatial Autocorrelation. Pion, London.
Cliff A.D., Ord J.K. 1981. Spatial Processes: Models and Applications. Pion, London.
Geary R. 1954. The Contiguity Ratio and Statistical Mapping. The Incorporated Statistician, 5: 115–145.
Getis A. 2007. Reflections on spatial autocorrelation. Regional Science and Urban Economics, 37: 491–496.
Getis A., Aldstadt J. 2004. Constructing the Spatial Weights Matrix Using a Local Statistic. Geographical Analysis, 36: 90–104.
Getis A., Ord J.K. 1992. The Analysis of Spatial Association by Use of Distance Statsitsics. Geographical Analysis, 24: 189–206.
Griffith D.A. 1996. Some Guidelines for Specifying the Geographic Weights Matrix Contained in Spatial Statistical Models. [W:] S.L. Arlinghaus (red.), Practical Handbook of Spatial Statistics. CRC, Boca Raton.
Griffith D.A. 2000. Eigenfunction properties and approximation of selected incidence matrices employed in spatial analysis. Linear Algebra and its Applications, 321: 95–112.
Griffith D.A. 2003. Spatial Autocorrelation and Spatial Filtering, Springer, Berlin–Heidelberg.
Griffith D.A. 2004. Extreme eigenfunctions of adjacency matrices for planar graphs employed in spatial analyses. Linear Algebra and its Applications, 388: 201–219.
Griffith D.A. 2010. The Moran coefficient for non-normal data. Journal of Statistical Planning and Inference, 140: 2980–2990.
Haining R.P. 1978. Estimating Spatial Interaction Models. Environment and Planning A, 10: 305–320.
Haining R.P. 1995. Data problems in spatial econometric modeling. [W:] L. Anselin, R.J.G.M. Florax (red.), New Directions in Spatial Econometrics. Springer-Verlag, Berlin, s. 156–171.
Hauke J., Kossowski T. 2008. Moran's coefficient and algebraic characteristics of some standardized connectivity matrices used to measure a spatial autocorrelation. Working Paper of The 17th International Workshop in Matrices and Statistics in honour of Professor Theodore Wilbur Anderson 90th birthday, July 23–26. Tomar, Portugal.
Janc K. 2006. Zjawisko autokorelacji przestrzennej na przykładzie statystyki I Morana oraz lokalnych wskaźników zależności przestrzennej (LISA) – wybrane zagadnienia metodyczne. [W:] T. Komornicki, Z. Podgórski (red.), Idee i praktyczny uniwersalizm geografii. Dokumentacja Geograficzna, 33: 76–83.
Janc K. 2007. Wpływ kapitału ludzkiego na efektywność gospodarek lokalnych w Polsce – przykład zastosowania regresji przestrzennej. [W:] P. Brezdeń, S. Grykień (red.), Regionalny wymiar integracji europejskiej. T. IX. IGiRR, Uniwersytet Wrocławski, Wrocław, s. 87–98.
Kelejian H.H., Prucha I. 1997. 2SLS and OLS in spatial autoregressive model with equal weights. Regional Science and Urban Economics, 32: 691–707.
Kelejian H.H., Prucha I. 1998. A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. Journal of Real Estate Finance and Economics, 17, 1: 99–121.
Kelejian H.H., Prucha I. 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review, 40, 2: 509–533.
Kelejian H.H., Tavlas G.S., Hondroyiannis G. 2006. A Spatial Modelling Approach to Contagion Among Emerging Economies. Open Economies Review, 17: 423–441.

Kopczewska K. 2006. Ekonometria i statystyka przestrzenna z wykorzystaniem programu R CRAN. Cedewu.pl, Warszawa.
Kosfeld R., Eckey H.-F., Dreger C. 2006. Regional Productivity and Income Convergence in the Unified Germany, 1992–2000. Regional Studies, 40, 7: 755–767.
Kossowski T. 2006. Modelowanie struktury sieci transportowej regionu wielkopolskiego. Bogucki Wydawnictwo Naukowe, Poznań.
Kossowski T. 2009. Metody i modele ekonometrii przestrzennej. [W:] Z. Zwoliński (red.), GIS – platforma integracyjna geografii. Bogucki Wydawnictwo Naukowe, Poznań, s. 145–165.
Kossowski T. 2009a. Konwergencja przestrzenna – aspekty teoretyczne. [W:] P. Churski (red.), Praktyczne aspekty badań regionalnych – varia. Vol. II. Biuletyn Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM, Seria Rozwój Regionalny i Polityka Regionalna 8: 7–20.
Kossowski T., Motek P. 2009. Spatial modelling of the local public finance in Poland. [W:] T. Markowski, M. Turała (red.), Theoretical and practical aspects of urban and regional development. Studia Regionalia 24: 152–167.
Lee L.-F. 2007. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. Journal of Econometrics, 137: 489–514.
Lee S.-I. 2001. Developing a bivariate spatial association measure: An integration of Pearson's r and Moran's I. Journal of Geographical Systems, 3: 369–385.
Maddala G.S. 2006. Ekonometria, Wydawnictwo Naukowe PWN, Warszawa.
Maćkiewicz B. 2007. Rynek nieruchomości niezabudowanych w Poznaniu i powiecie poznańskim w latach 1995–2000. Bogucki Wydawnictwo Naukowe, Poznań.
Mobley L., Root E., Anselin L., Lozano N.. Koshinsky J. 2006. Spatial analysis of elderly access to primary care services. International Journal of Health Geographics, 5: 19.
Moran P.A.P. 1950. Notes on continuous stochastic phenomena. Biometrika, 37: 17–23.
Olejnik A. 2008. Using the spatial autoregressively distributed lag model in assessing the regional convergence of per-capita income in the EU25. Papers in Regional Science, 87, 3: 371–384.
Oud J.H.L., Folmer H., Patuelli R., Nijkamp P. 2010. A Spatial-Dependence Continuous-Time Model for Regional Uemployment in Germany. 50th ERSA Congress in Jonkoping, Working Paper.
Paelinck J.H.P. 2000. On aggregation in spatial econometric modeling. Journal of Geographical Systems, 2: 157–165.
Paelinck J.H.P., Klaassen L.H. 1979. Spatial Econometrics. Gower, Westmead, Farnborough.
Ratajczak W. 1980. Analiza i modele wpływu czynników społeczno-gospodarczych na kształtowanie się sieci transportowych. PWN, Warszawa–Poznań.
Ratajczak W. 2008. Modele ekonometrii przestrzennej w analizie regionalnej. [W:] T. Stryjakiewicz, T. Czyż (red.), O nowy kształt badań w geografii i gospodarce przestrzennej, Biuletyn KPZK PAN, 237: 186–202.
Suchecki B. (red.) 2010. Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych. Wydawnictwo C.H. Beck, Warszawa.
Tiefelsdorf M., Boots B. 1995. The exact distribution of Moran's I. Environment and Planning A, 27: 985–999.
Tiefelsdorf M., Griffith D.A., Boots B. 1998. A Variance Stabilizing Coding Scheme for Spatial Link Matrices. Environment and Planning A, 31: 165–180.
Tobler W. 1970. A computer model simulating urban growth in Detroit region. Economic Geography, 46, 2: 234–240.
Vazquez E.F. 2010. Empirical versus exogenous spatial weighting matrices: an entropy-based intermediate solution. 50th ERSA Congress in Jonkoping, Working Paper.

Walde J., Larch M., Tappeiner G. 2008. Performance contest between MLE and GMM for huge spatial autoregressive models. Journal of Statistical Computation and Simulation, 78, 2: 151–166.
Ying L.G. 2003. Understanding China's recent growth experience: A spatial econometric perspective. The Annals of Regional Science, 37, 4: 613–628.
Zelias A. (red.) 1991. Ekonometria przestrzenna. PWE, Warszawa.

Theoretical aspects of spatial modeling in regional research

Abstract: Spatial modeling is currently one of the primary research tools used in regional analysis. Spatial models are an extension of traditional econometric models, which are included in the so-called spatial effects: spatial dependence and spatial heterogeneity. The article presents the theoretical basis of spatial modelling, together with definitions of basic concepts and an analysis of their properties. Methods for estimating spatial models and diagnostics are presented. The study also indicates the complexity of spatial modeling, and the usefulness of this kind research approach. In this paper an outline the development trends of spatial modeling is delivered.

Keywords: spatial modeling, spatial dependence, spatial heterogeneity, weight matrix, estimation of spatial models