

Przestrzenne modelowanie wypadków drogowych we Włoszech

Maryna Abdurakhmanava
72720

19 września 2018

Abstrakt

Celem pracy jest omówienie wyników przestrzennego modelowania bezpieczeństwa ruchu drogowego we Włoszech ocenianego liczbą wypadków mających miejsce na drogach publicznych, w strefach zamieszkania lub strefach ruchu. Przeanalizowano wpływ udziału młodych mężczyzn (18-24 lat), gęstości zaludnienia oraz liczby środków transportu samochodowego na mieszkańca na liczbę wypadków drogowych. Wyniki wskazują na istnienie zjawisk przestrzennych, które oddziałują na liczbę wypadków drogowych w poszczególnych prowincjach. Wśród przetestowanych form zależności przestrzennych i macierzy najlepsze wyniki otrzymano w modelu SDM zbudowany przy użyciu macierzy sąsiedztwa oparte na odwrotności kwadratu odległości między centroidami do 200km z kryterium wygasania 2.

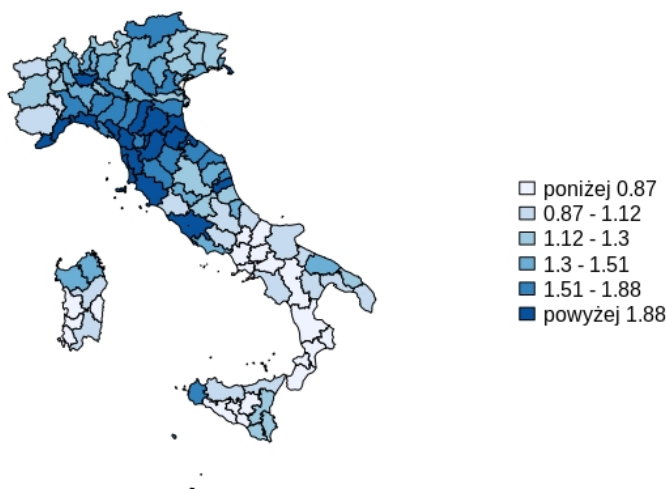
1 Wstęp

W 2015 roku policji we Włoszech zgłoszono 174 539 wypadków drogowych, w wyniku których uczestnik ruchu drogowego został ranny (171 303) lub doszło do jego śmierci (3 236). Ta kwestia często poruszana jest w pracach dotyczących ekonometrii przestrzennej. Zdecydowana większość analiz wskazuje, że głównym czynnikiem wypadków samochodowych są niebezpieczne zachowania poszczególnych grup użytkowników dróg, przede wszystkim jazda z nadmierną prędkością lub niedostosowania prędkości do warunków ruchu, niski stopień stosowania urządzeń bezpieczeństwa oraz nietrzeźwość uczestników ruchu drogowego. Szczególną uwagę zwraca się na grupę młodych kierowców w wieku 18-24 lata. Jest to grupa osób, którą cechuje brak doświadczenia w kierowaniu pojazdami oraz skłonność do zachowań ryzykownych. Na występowanie wypadków drogowych wpływ ma natężenie ruchu, a także warunki atmosferyczne i oświetlenie, przy czym ten ostatni czynnik jest uzależniony od pory dnia i pory roku. Znaczącą grupę ofiar i w mniejszym stopniu sprawców wypadków drogowych stanowią rowerzyści i motocykliści.

2 Zmienna zależna

Główną zmienną analizowaną w niniejszej pracy jest liczba wypadków drogowych w przeliczeniu na tysiąc osób, których skutkiem jest śmierć jednego z uczestników wypadku lub obrażenia ciała. Wypadki te miały miejsce na drogach publicznych, w strefach zamieszkania lub strefach ruchu. Obserwacje pochodzą z danych włoskiego głównego urzędu statystycznego (*ISTAT*) z 2015 roku. Przedmiotowa zmienna została zaprezentowana na mapie na rysunku 1. Najbardziej widoczne są skupienia w centralo- i północno-zachodniej części kraju (o wyższych wartościach zmiennej) oraz w części południowej – o niższych poziomach zmiennej zależnej.

Rysunek 1: Wykres zmiennej liczba wypadków drogowych na tysiąc osób



3 Zmienne niezależne

Pierwszą zastosowaną zmienną objaśniającą jest odsetek młodych mężczyzn (18-24 lat). Obserwacje te również pochodzą z bazy włoskiego głównego urzędu statystycznego (*ISTAT*). Dla tej zmiennej można zaobserwować wyraźny podział na część północną i centralno-północną (o niższych poziomach zmiennej) a część południową centralno-południową (o wyższych wartościach zmiennej).

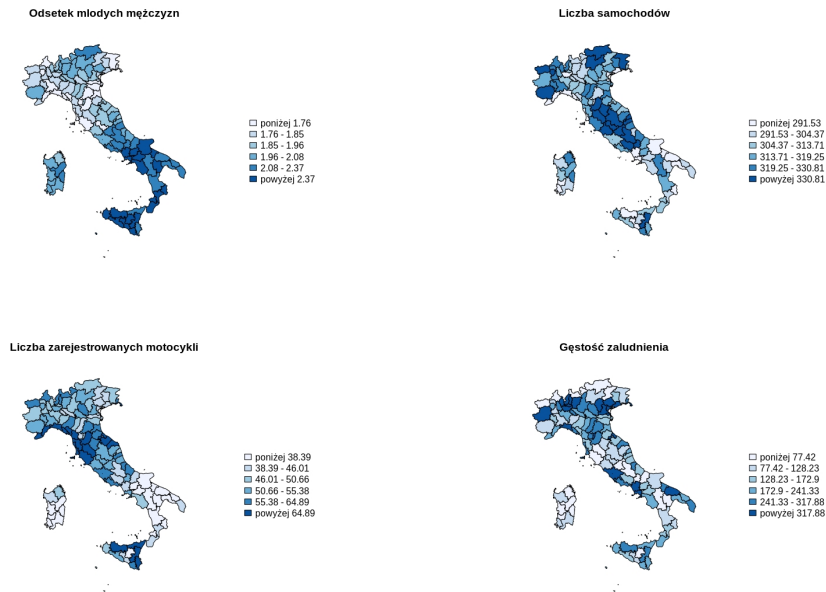
Jako kolejne zmienne wykorzystano dane dotyczące liczby samochodów oraz liczby motocykli w przeliczeniu na tysiąc osób. Obserwacje pochodzą z bazy (*ISTAT*). Dla tych zmiennych można zaobserwować kilka skupisk: największe w części centralnej (liczba samochodów) i północno-zachodniej (liczba motocykli), mniejsze w części północnej oraz wschodniej części

Sycylii.

Ostatnia zmienna - gęstość zaludnienia - pochodzi ze zbiorów Urzędu Statystycznego Wspólnot Europejskich (EUROSTAT). Widocznych jest kilka niewielkich skupień o wyższych wartościach zmiennej.

Wyżej wymienione zmienne zostały zaprezentowane na rysunku 2.

Rysunek 2: Wykresy zmiennych niezależnych



4 Macierze W

Zasadniczym elementem analizy przestrzennej jest określenie struktury przestrzennego sąsiedztwa. W przeprowadzonych badaniach przyjęto 7 sposobów definiowania macierzy wag:

1. W1 - Sąsiedztwo I rzędu (macierz oparta na kryterium wspólnej granicy i normalizowana wierszami).
2. W2 - Sąsiedztwo II rzędu.
3. W3 - Sąsiedztwo I rzędu (macierz oparta na kryterium wspólnej granicy i normalizowana największą wartością własną).
4. W4 - Sąsiedztwo oparte na odwrotności kwadratu odległości między centroidami (do 200 km z kryterium wygasania równym 2).
5. W5 - Sąsiedztwo o promieniu 76 km.

6. W6 - 5-najbliższych sąsiadów.
7. W7 - Sąsiedztwo oparte na odległościach euklidesowych między 3 zmiennymi (stopa bezrobocia, współczynnik umieralności oraz udział liczby cudzoziemców w ogólnej populacji).

Wizualna analiza układu przestrzennego liczby wypadków drogowych na rysunku 1 wskazuje na tworzenie się skupisk niskiej jak i wysokiej liczby wypadków drogowych, co może świadczyć o silnych zależnościach przestrzennych. Dla potwierdzenia zostały przeprowadzone globalne testy Morana I, Gearyego C i lokalne testy Morana I i Gedisa-Orda dla zmiennej zależnej dla poszczególnych macierzy wag przestrzennych. Wyniki testów zostały przedstawione w tabelicy 1. Warto jednak zauważyć, że zgodnie z twierdzeniem Gaussa-Markowa, zastosowanie wyżej wymienionych testów do zmiennej zależnej jest w zasadzie nie konkluzywne, dlatego że nawet jeżeli proces istnieje może on zostać wyjaśniony przez model liniowy.

Tablica 1: Wyniki Globalnych testów Morana I, Gearyego C oraz testu liczby połączeń

	Morana I	Gearyego C	join-count (>1)	join-count (>1.20)	join-count (>1.50)
W1	9.1263***	8.6461***	5.471***	7.5109***	6.5193***
W2	13.077***	12.056***	7.2029***	11.08***	9.171***
W3	9.3451***	8.3702***	5.3925***	7.6214***	6.9336***
W4	11.475***	10.852***	6.4875***	10.418***	9.1263***
W5	8.8208***	8.5121***	6.2565***	7.6513***	9.1263***
W6	10.707***	9.7719***	5.3264***	8.7552***	8.1315***
W7	18.454***	13.423***	9.4569***	14.205***	8.5676***

* p-value '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Dla wszystkich macierzy dodatnie i wysokie (istotne statystycznie) wartości statystyki globalnej Morana I i Gearyego C potwierdziły słuszność wizualnej analizy układu przestrzennego liczby wypadków drogowych: na dowolnym poziomie istotności możemy odrzucić hipotezę zerową o braku globalnej autokorelacji przestrzennej.

Wyniki zastosowania lokalnej statystyki Morana dla poszczególnych macierzy wag przestrzennych wykazują istnienie lokalnych klastrów jak i jednostki, które nie uczestniczą w globalnym procesie przestrzennym.

W celu przeprowadzenia testu liczby połączeń, zmienna została podzielona na dwie części według wybranego progu: 1.1, 1.2 oraz 1.5. Statystyki testu przedstawione w tabelicy 1 wykazują możliwość odrzucenia hipotezy zerowej o losowym rozkładzie przestrzennym otrzymanych grup.

Na podstawie danych dokonano estymacji modelu liniowego. Wyniki przedstawione zostały w tabelicy 2.

Tablica 2: Model liniowy

	Estimate *	p-value
<i>intercept</i>	2.59439717	0.000000***
<i>car_per_pop</i>	-0.00143961	0.0796.
<i>motocycle_per_pop</i>	0.01421332	0.000000***
<i>young_perc</i>	-0.77520207	0.000000***
<i>density</i>	0.00014367	0.0600.
<i>F - statistic</i>	62.21	0.0000
<i>R - squared</i>	0.7033	
<i>AIC</i>	44.2895	

* p-value '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Zarówno dla odsetka młodych mężczyzn, jak i liczby samochodów w przeliczeniu na tysiąc osób otrzymano ujemną ocenę parametru mającym wpływ na zwiększenie poziomu tych zmiennych na spadek liczby wypadków drogowych, co nie jest wynikiem oczekiwanym. Zmienna liczba samochodów nie jest istotna statystycznie na przyjętym poziomie istotności 0.05.

Zmienna liczba motocykli oraz gęstość zaludnienia mają wpływ dodatni na liczbę wypadków drogowych, jak było oczekiwane, zmienna gęstość zaludnienia nie jest istotna statystycznie na przyjętym w analizie poziomie istotności.

Statystyka F jest bardzo wysoka, co pozwala odrzucić hipotezę zerową o łącznej nieistotności wszystkich zmiennych w modelu. R-square jest relatywnie duży: ten model wyjaśnia około 70% zmienności zmiennej objaśnianej. Kryterium informacyjny Akaike zostanie użyty do porównania modelu liniowego, z modelami przestrzennymi.

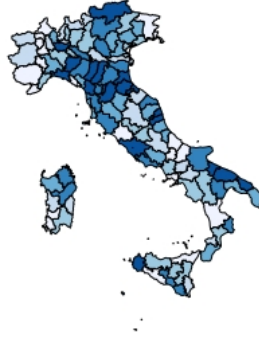
Wykres reszt modelu został przedstawiony poniżej na rysunku 3. Najbardziej widoczne są skupienia w części centralno- i północno-zachodniej, więc istnienie procesów przestrzennych w resztach wskazuje na to, że niezbędne jest przejście na modelowanie przestrzenne.

Jednak przed przejściem do modelowania przestrzennego, przeprowadzono globalne testy Morana I, Gearyego C i lokalne testy Morana I i Gedisa-Orda dla reszt modelu liniowego dla poszczególnych macierzy wag przestrzennych. Wyniki testów zostały przedstawione w tablicy 3.

Dla wszystkich macierzy dodatnie i wysokie (istotne statystycznie) wartości statystyki globalnej Morana I i Gearyego C potwierdziły słuszność wizualnej analizy układu przestrzennego reszt modelu KMNK: na dowolnym poziomie istotności możemy odrzucić hipotezę zerową o braku globalnej przestrzennej autokorelacji reszt z regresji liniowej. Taki wynik wskazuje na to, że proces jest nieprawidłowo wyspecyfikowany, czyli że liniowa specyfikacja nie jest tą, która prawidłowo odzwierciedla proces generujący dane. W tym procesie dane są powiązane przestrzennie i te powiązania również muszą być brane pod uwagę.

Wyniki zastosowania lokalnej statystyki Morana dla poszczególnych macierzy wag przestrzen-

Rysunek 3: Wykres reszt dla modelu liniowego



Tablica 3: Wyniki Globalnych testów Morana I, Gearyego C oraz testu liczby połączeń

	Morana I	Gearyego C	join-count (>0)	join-count (>-0.1)	join-count (>0.1)
W1	4.8117***	4.7573***	4.0722***	2.3016***	2.7444***
W2	4.8837***	4.1584***	2.9219***	2.3266***	2.5503***
W3	5.0399***	4.6375***	3.6994***	2.3943***	3.0905***
W4	5.8285***	5.1077***	4.6642***	2.8477***	2.7395***
W5	4.8224***	4.7346***	4.6642***	3.1746***	2.8869***
W6	5.0304***	4.6902***	4.6562***	3.2751***	3.1911***
W7	5.0945***	3.8784***	1.0596.	2.2819***	0.80609

* p-value '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

nych wskazują na to, że istnieją lokalne klastry, jak i jednostki, które nie uczestniczą w globalnym procesie przestrzennym.

W celu przeprowadzenia testu liczby połączeń, reszty z modelu liniowego zostały podzielone na dwie części według wybranego progu: 0,0,1,-0,1. Statystyki testu przedstawione w tablicy 3 wskazują na to, że możemy odrzucić hipotezę zerową o losowym rozkładzie przestrzennym otrzymanych grup dla macierzy W1-W6. Dla macierzy W7 nie ma podstaw do odrzucenia hipotezy zerowej, gdy próg jest 0 oraz 0.1, jednak dla progu mniejszego niż 0 możemy odrzucić hipotezę zerową.

5 Modele z jednym źródłem oddziaływania przestrzennego

W związku z tym, że wyniki testowania obecności procesów przestrzennych niewiele różnią się dla poszczególnych macierzy wag przestrzennych dla dalszej analizy zostały wybrane trzy macierze różniące się według metody ich konstrukcji :

1. W1 - Sąsiedztwo I rzędu (macierz oparta na kryterium wspólnej granicy i normalizowana wierszami).
2. W4 - Sąsiedztwo oparte na odwrotności kwadratu odległości między centroidami (do 200 km z kryterium wygasania równym 2).
3. W6 - 5-najbliższych sąsiadów.

Na bazie przyjętych macierzy wag przestrzennych przeprowadzona została powtórna analiza przestrzenna liczby wypadków drogowych w przeliczeniu na tysiąc osób we Włoszech w 2015 roku mierzonej w układzie powiatów i dla modelu liniowego zostały przeprowadzone testy LM i odporne testy LM na bazie wybranych macierzy wag przestrzennych. Wyniki testów LMerr oraz LMlag pozwalają nam odrzucić hipotezę zerową o tym, że model jest liniowy na rzecz hipotezy alternatywnej, że model jest modelem SME (test LMerr) lub SLM (test LMlag). To świadczy o obecności procesu przestrzennego, jednak żaden z tych testów nie jest konkluzyjny ponieważ ogranicza wybór tylko dwoma modelami. Dla macierzy W1 oba testy odporne wskazują na brak procesu przestrzennego, i dla każdej z macierzy nie mamy podstaw do odrzucenia hipotezy zerowej o braku opóźnienia zmiennej objaśnianej. Dla macierzy W4 i W6 są podstawy do odrzucenia hipotezy zerowej o braku autokorelacji przestrzennej.

Tablica 4: Wyniki testów LM dla modelu liniowego

	W1	W4	W6
<i>LMerr</i>	17.891***	24.756***	18.355***
<i>LMlag</i>	17.448**	21.305***	17.195***
<i>RLMerr</i>	2.5342	5.8138*	3.6071*
<i>RLMlag</i>	2.0914	1 2.3637	2.4467

* p-value ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘,’ 1

5.1 Model pure SAR

Na podstawie danych dokonano estymacji trzech przestrzennych modeli czystej autoregresji przestrzennej (pure SAR), w zależności od przyjętej macierzy wag. Uzyskane wyniki przedstawione zostały w tablicy 5. Współczynnik autokorelacji przestrzennej rho w każdej z tych modeli jest dodatni i istotny statystycznie, co wskazuje na istnienie autokorelacji przestrzennej. Najniższe kryterium Akaike ma model zbudowany z wykorzystaniem macierzy W4.

Tablica 5: Model pure SAR

	W1	W4	W6
<i>intercept</i>	1.40254***	1.42735***	1.41687***
ρ	0.70763 * **	0.78649***	0.72238***
<i>LR test</i>	63.427***	68.754***	65.287***
<i>AIC</i>	108.5	103.17	106.64
<i>Moran</i>	-0.57643	0.032302	-0.38185

* p-value '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.2 Model SAR

Przy przejściu do modelu SAR zostają uwzględnione regresory. Wyniki przedstawione są w tablicy 6 i wskazują na zasadność użycia procesu przestrzennego dla wyjaśnienia badanego zjawiska - parametr ρ jest dodatni i istotny statystycznie (również statystyki LR oraz Wald wskazują na możliwość odrzucenia hipotezy zerowej o nieistotności ρ).

Tablica 6: Model SAR

	W1	W4	W6
<i>intercept</i>	1.434876462 ***	1.337457801***	1.494828770***
<i>car_per_pop</i>	-0.001301660.	-0.001590192*	-0.001402687.
<i>motocycle_per_pop</i>	0.012374151***	0.012308584 ***	0.012409677***
<i>young</i>	-0.422458596***	-0.369954551***	-0.428215759 ***
<i>density</i>	0.000153632*	0.000135648*	0.000150308*
ρ	0.36654 * **	0.43343***	0.35475**
<i>LR test</i>	16.056 * **	17.42***	14.168***
<i>Wald statistic</i>	18.504 * **	20.513***	15.541***
<i>AIC</i>	30.233	28.869	32.121
<i>Moran I global</i>	1.0751	1.5096.	1.1529

* p-value '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Reszty z modelu SAR zbudowanego na podstawie macierz wag przestrzennych W1 oraz W6 nie wykazują przestrzennej autokorelacji, na co wskazują testy Morana. Kryterium Akaike jest o wiele niższe niż dla modeli pure SAR, a najniższy poziom wśród modeli SAR ma model

zbudowany przy użyciu macierzy W4. Wszystkie wybrane do modelu zmienne niezależne są istotne na poziomie istotności 0.1 dla macierzy W4 i 0.05 dla macierzy W1 oraz W6.

5.3 Model SEM

Następnie zbudowano model autoregresji przestrzennej składnika losowego SEM. Wyniki przedstawione w tablicy 7 wskazują na zasadność użycia procesu przestrzennego dla wyjaśnienia badanego zjawiska - parametr lambda jest dodatni i istotny statystycznie (na co również wskazują test LR oraz statystyka Walda). Reszty ze wszystkich modeli SEM nie wykazują przestrzennej autokorelacji, na co wskazują testy Morana. Kryterium Akaike jest nieco niższe niż dla modeli SAR, a najniższy wśród modeli SEM znowu ma model zbudowany przy użyciu macierzy W4.

Tablica 7: Model SEM

	W1	W4	W6
<i>intercept</i>	1.799615692***	1.820763376***	1.774111622***
<i>car_per_pop</i>	-0.001005634	-0.001227228.	-0.000988635
<i>motocycle_per_pop</i>	0.015196503***	0.014662237***	0.014660919***
<i>young</i>	-0.482536886***	-0.440373873*	-0.458875399***
<i>density</i>	0.000178048*	0.000156982*	0.000183187**
λ	0.53789***	0.6565***	0.55951***
<i>LR test</i>	20.8 * **	21.078***	16.863***
<i>Wald statistics</i>	34.961 * **	49.123***	33.901***
<i>AIC</i>	27.751	25.212	29.427
<i>Moran I global</i>	-0.18792	0.26571	-0.20541

* p-value '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4 Model SLX

Wyniki ostatniego z prostych modeli - modelu SLX - zostały przedstawione w tablicy 8. Reszty ze wszystkich modeli SLX wykazują przestrzenną autokorelację, na co wskazują testy Morana. Warto jednak zauważyć, że parametry przy opóźnieniach zmiennej *young* są istotne statystycznie na poziomie istotności 0.05 dla wszystkich macierzy wag przestrzennych.

Tablica 8: Model SLX

	W1	W4	W6
<i>intercept</i>	3.42011292 ***	3.72886786***	3.81268892***
<i>car_per_pop</i>	-0.00154420.	-0.00155271.	-0.00153795.
<i>motocycle_per_pop</i>	0.01458977***	0.01334377***	0.01395772***
<i>young</i>	-0.01617179	-0.22034072	-0.07432654
<i>density</i>	0.00015446.	0.00016839.	0.00015013.
<i>lag.car_per_pop</i>	-0.00143151	-0.00263972	-0.00242082
<i>lag.motocycle_per_pop</i>	-0.00040797	0.00291989	0.00007329
<i>lag.young</i>	-0.90059599**	-0.70297548*	-0.87623021**
<i>lag.density</i>	-0.00022176	-0.00031537	-0.00021786 .
<i>R – squared</i>	0.732	0.7232	0.7338
<i>F – statistic</i>	34.48 * **	32.99***	34.81***
<i>AIC</i>	41.078	44.6357	40.3314
<i>Moran I global</i>	4.8915***	5.692***	4.8884***

* p-value ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

5.5 Porównanie modeli z jednym źródłem oddziaływania przestrzennego

Na tym etapie z analizy zostają wyłączone 2 modele pure SAR oraz SLX. Model pure SAR używany jest w ograniczonej klasie sytuacji, gdy jedyną przyczyną są wartości tej samej zmiennej w sąsiedztwie, jednak przy modelowaniu bezpieczeństwa ruchu drogowego nie mamy doczynienia z podobną sytuacją, na co również wskazują wyniki modelu liniowego. Do tego kryterium Akaike jest o wiele wyższy niż dla pozostałych modeli.

Jak już było wspomniane wcześniej, reszty ze wszystkich specyfikacji modelu SLX wykazują przestrzenną autokorelację.

Dla porównania modeli SEM i SAR został użyty test likelihood ratio (LR) dla poszczególnych macierzy. Wynik jest przedstawiony w tabeli 9 i wskazuje na to że nie mamy podstaw do odrzucenia hipotezy zerowej o tym, że model SAR pokazuje lepsze wyniki, niż model SEM. Na tym etapie najlepszym modelem wydaje się być model SAR oparty o macierz W4 (sąsiedztwo oparte na odwrotności kwadratu odległości między centroidami do 200km z kryterium wygasania 2), na który wskazują najniższy kryterium Akaike oraz istotności statystyczne na poziomie istotności 0.05 wszystkich zmiennych niezależnych dobranych do modelu.

Tablica 9: Porównanie modeli SAR i SEM

	W1	W4	W6
<i>LR</i>	0.289022	0.1606495	0.2599228

6 Modele z dwoma źródłami oddziaływania przestrzennego

6.1 Model SARAR

Pierwszym wyestymowanym modelem z dwoma źródłami oddziaływania przestrzennego, jest model SARAR, łączący w sobie dwa źródła procesu przestrzennego, zarówno autokorelacje zmiennej objaśnianej, tak i autokorelacje reszt. Na podstawie wyników zaprezentowanych w tablicy 10, możemy wskazać, że parametr ρ jest nieistotny statystycznie dla wszystkich specyfikacji modelu SARAR, a parametr λ jest istotny statystycznie na każdym poziomie istotności dla macierzy wag przestrzennych W3 oraz W4, więc możemy uprościć ten model do modelu z jednym źródłem oddziaływania przestrzennego - SEM. Reszty ze wszystkich modeli SARAR nie wykazują przestrzennej autokorelacji, na co wskazują testy Morana.

Tablica 10: Model SARAR

	W1	W4	W6
<i>intercept</i>	1.668911956***	1.651320200***	1.704019245***
<i>car_per_pop</i>	-0.001081679	-0.001265726.	-0.000896396
<i>motorcycle_per_pop</i>	0.0152623710.014630549***	0.014282116***	0.014829839***
<i>young</i>	-0.482649030***	-0.446689092*	-0.16384676
<i>density</i>	0.000178161*	0.000161198*	0.000165426***
ρ	0.13774	0.15897	-0.43149
λ	0.4002***	0.51758***	0.83611.
<i>LR test</i>	18.996***	21.444***	17.456***
<i>AIC</i>	29.293	26.845	30.834
<i>Moran I global</i>	-0.13865	0.25105	-0.24383

* p-value '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

6.2 Model SDM

W dalszej kolejności w tabeli 11 został przedstawiony model SDM, łączący w sobie dwa proste modele - SAR oraz SLX. Istotnymi okazują się być parametry przy opóźnionej zmiennej liczba motocykli dla wszystkich specyfikacji oraz parametry przy opóźnionych zmiennych liczba młodych mężczyzn oraz gęstość zaludnienia. Parametr ρ okazuje się być istotny statystycznie dla wszystkich specyfikacji modeli. Reszty ze wszystkich modeli SDM nie wykazują przestrzennej autokorelacji, na co wskazują testy Morana.

Tablica 11: Model SDM

	W1	W4	W6
<i>intercept</i>	1.7197581069***	1.58869306.	2.044398533***
<i>car_per_pop</i>	-0.0013026002.	-0.00142773.	-0.001338897.
<i>motocycle_per_pop</i>	0.015551567***	0.01448208***	0.014724478***
<i>young</i>	0.0670160599	-0.14632742	0.014605473
<i>density</i>	0.0001594409*	0.00015407*	0.000156489*
<i>lag.car_per_pop</i>	0.0000071665	-0.00016460	-0.000611514
<i>lag.motocycle_per_pop</i>	-0.0083637552***	-0.00866066*	-0.008324817***
<i>lag.young</i>	-0.5419329579*	-0.25113118	-0.521337850.
<i>lag.density</i>	-0.0002148201	-0.00022840	-0.000229385*
ρ	0.47203***	0.5797***	0.4697***
<i>AIC</i>	25.916	28.639	28.207
<i>Moran I global</i>	-0.048386	0.40064	0.0091069

* p-value '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

6.3 Model SDEM

Ostatnim z modeli z dwoma źródłami oddziaływania przestrzennego został oszacowany model SDEM, który jest kombinacją modelu SEM (z autokorelacją reszt) z modelem SLX (z opóźnieniem przestrzennym zmiennych objaśniających). Na podstawie wyników zaprezentowanych w tablicy 12, możemy wskazać, że parametr λ jest istotny statystycznie dla wszystkich specyfikacji modelu SDEM oraz istotnymi okazują się być parametry przy opóźnionej zmiennej liczba młodych mężczyzn dla wszystkich specyfikacji modelu. Kryterium Akaike okazują się być jedne z najniższe wśród wszystkich przebadanych modeli.

Tablica 12: Model SDEM

	W1	W4	W6
<i>intercept</i>	2.773957459***	2.9399918789*	3.436141880***
<i>car_per_pop</i>	-0.001296185.	-0.0014362767*	-0.001401262.
<i>motocycle_per_pop</i>	0.015530832***	0.0150480281***	0.014942330 ***
<i>young</i>	-0.004255788	-0.1433614564	-0.031960314
<i>density</i>	0.000148213*	0.0001352927.	0.000142878*
<i>lag.car_per_pop</i>	0.000151353	0.0004572118	-0.001217445
<i>lag.motocycle_per_pop</i>	-0.000097244	-0.0049893785	-0.003889153
<i>lag.young</i>	-0.838537041***	-0.7667121623*	-0.871883424***
<i>lag.density</i>	-0.000131301	0.0000083853	-0.000168507
λ	0.48011***	0.62707***	0.48425***
<i>AIC</i>	26.24	27.746	28.778
<i>Moran I global</i>	-0.044808	0.34843	0.054722

* p-value ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

6.4 Wybór najlepszego modelu

Z dalszej analizy został wykluczony model SARAR, ponieważ parametr ρ okazuje się być nieistotny i model upraszcza się do modelu SEM. W tablicy 13 zostali przedstawione wyniki porównań modeli SAR, SEM, SDM oraz SDEM. Dla porównania modeli został użyty test likelihood ratio (LR) dla poszczególnych macierzy. Testy LR wskazują, że modele z dwoma

Tablica 13: Porównanie modeli SAR, SEM, SDM, SDEM

	W1	W4	W6
<i>SAR vs SEM</i>	0.289022	0.1606495	0.2599228
<i>SAR vs SDM</i>	0.002115161	0.01632726	0.002587434
<i>SAR vs SDEM</i>	0.002487356	0.01044721	0.003442156
<i>SEM vs SDM</i>	0.007318337	0.1016328	0.009954626
<i>SEM vs SDEM</i>	0.008606114	0.06503105	0.01324299
<i>SDM vs SDEM</i>	1	0.6398627	1

źródłami procesu przestrzennego lepiej objaśniają zmienną, jednak dla modeli z dwoma źródłami procesu przestrzennego zdecydowanie lepiej prezentuje się model SDM. Najlepszym modelem wydaje się być model SDM oparty o macierz W4 (sąsiedztwo oparte na odwrotności kwadratu odległości między centroidami do 200km z kryterium wygasania 2).

7 Efekty przestrzenne

Dla najlepszej specyfikacji modelu SDM zostały wyznaczone średnie efekty bezpośrednie, pośrednie oraz całkowite. Efekty pośrednie okazały się być istotne na każdym poziomie

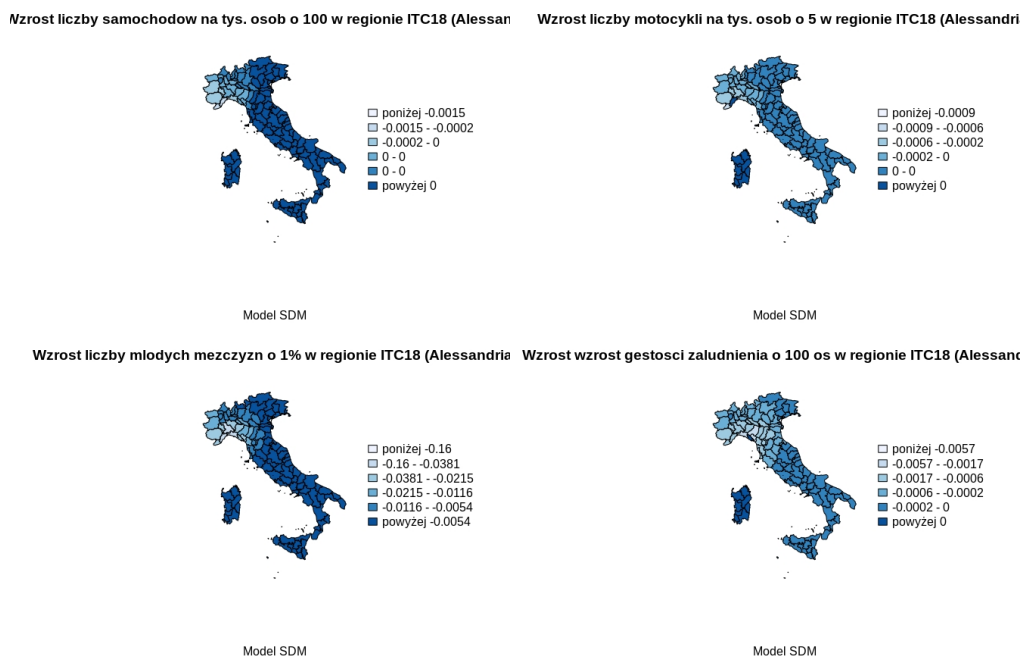
Tablica 14: Efekty przestrzenne dla modelu SDM opartym na W4

	Efekt bezpośredni	Efekt pośredni	Efekt całkowity
<i>car_per_pop</i>	-0.0015244	-0.0022410	-0.0037654
<i>motorcycle_per_pop</i>	0.0144508	-0.0007561	0.0136947
<i>young_per</i>	-0.1862653	-0.7861993	-0.9724646
<i>density</i>	0.0001401	-0.0003383	-0.0001982

istotności dla wszystkich zmiennych w modelu.

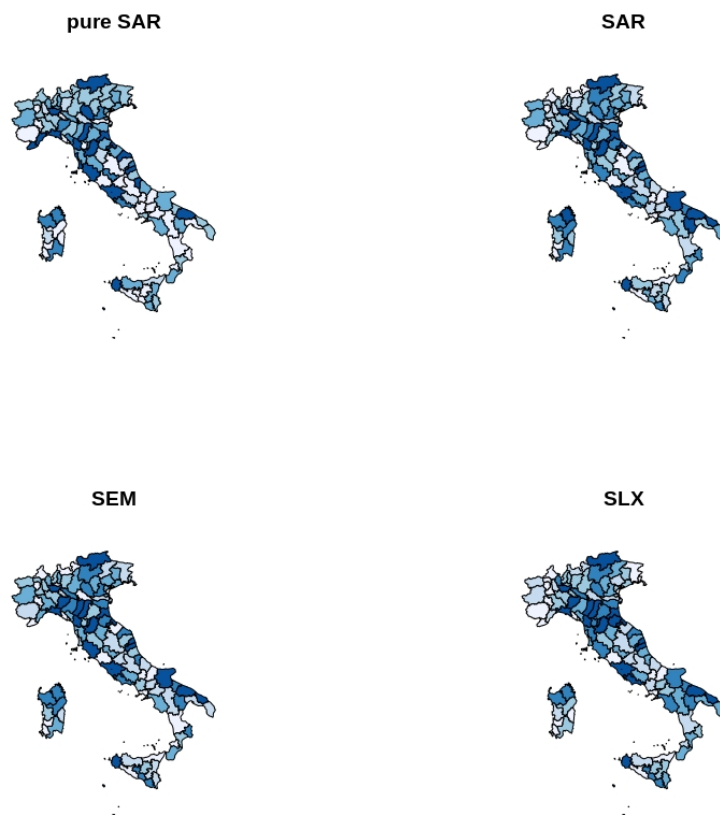
Na rysunku 4 przedstawiono efekty w poszczególnych regionach przy zadaniu impulsu w prowincji ITC18.

Rysunek 4: Efekt wpływu impulsu w prowincji ITC18 na poszczególne regiony

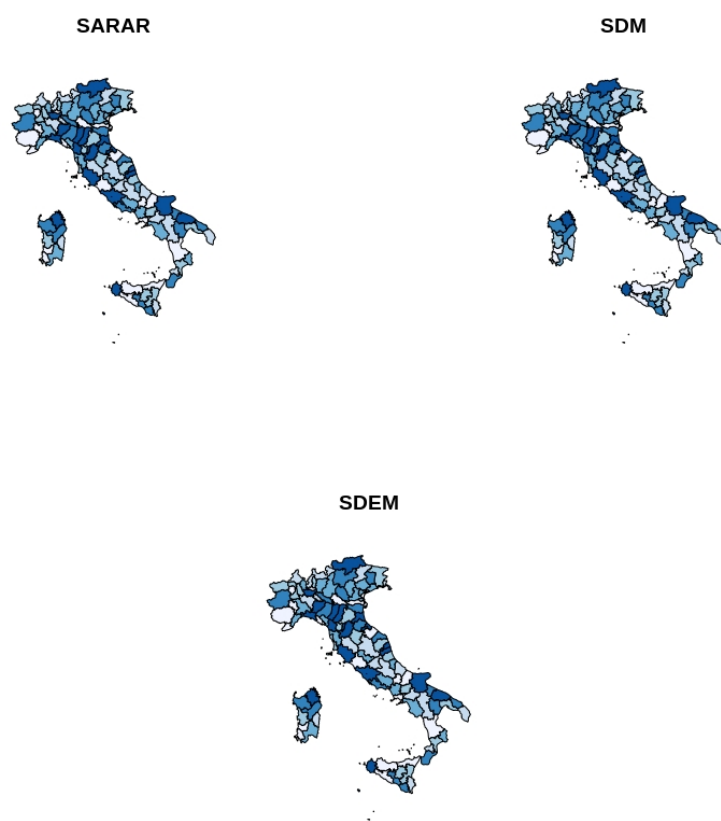


8 Wykresy reszt dla modeli zbudowanych w oparciu o macierz W1, W4, W6

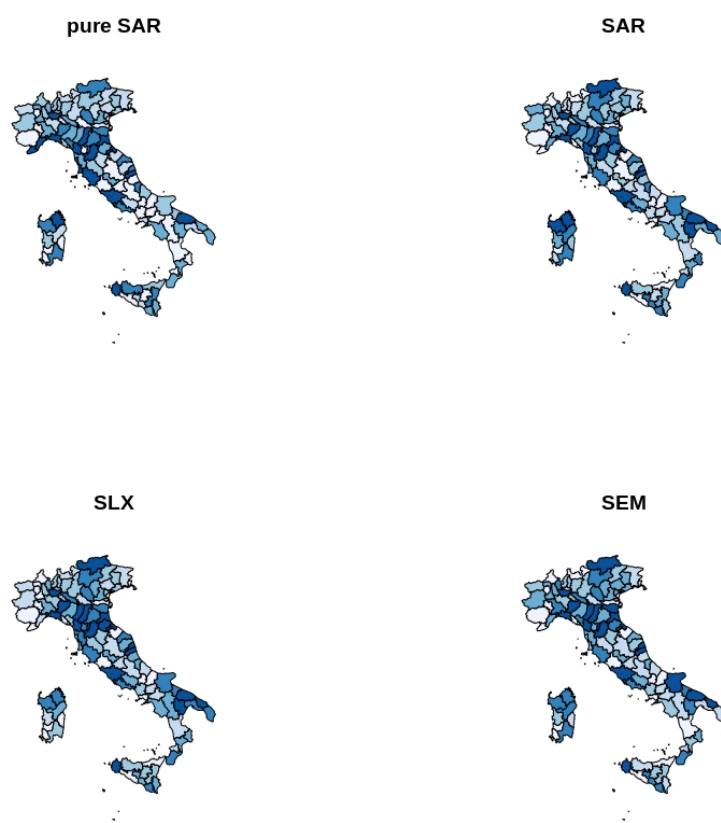
Rysunek 5: Wykresy reszt dla modeli zbudowanych w oparciu o macierz W1



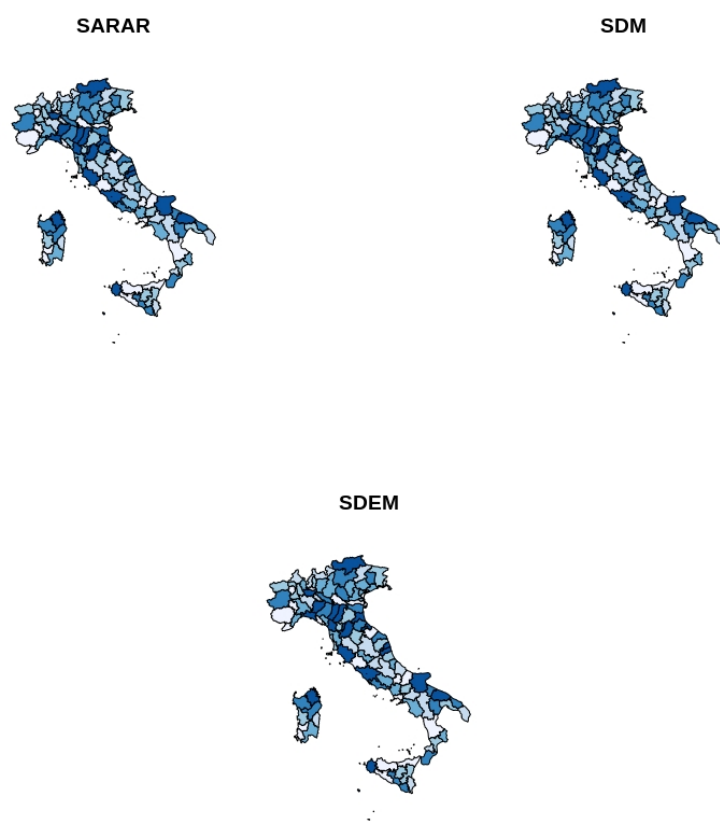
Rysunek 6: Wykresy reszt dla modeli zbudowanych w oparciu o macierz W_1



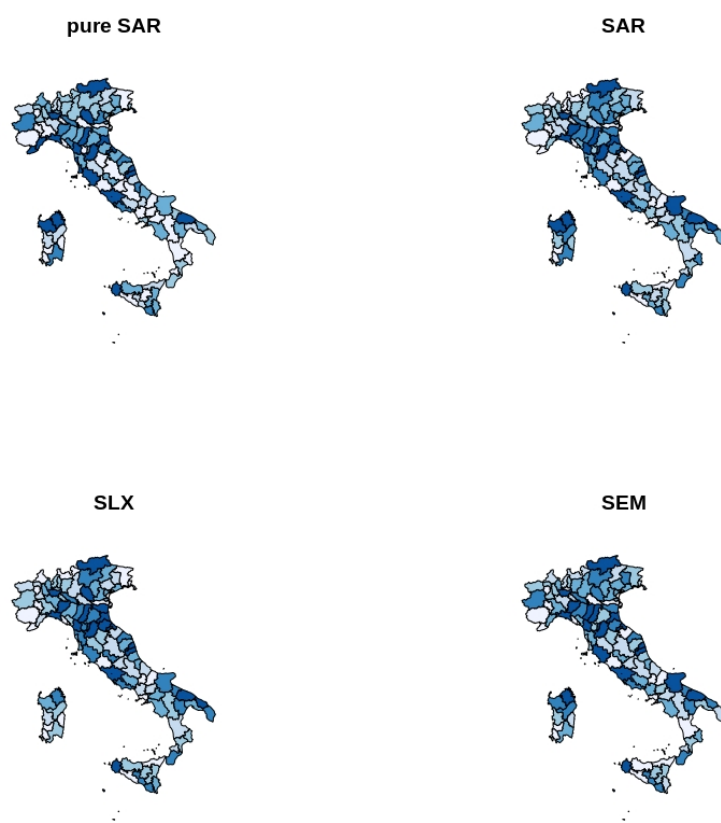
Rysunek 7: Wykresy reszt dla modeli zbudowanych w oparciu o macierz W_4



Rysunek 8: Wykresy reszt dla modeli zbudowanych w oparciu o macierz W_4



Rysunek 9: Wykresy reszt dla modeli zbudowanych w oparciu o macierz W6



Rysunek 10: Wykresy reszt dla modeli zbudowanych w oparciu o macierz W_6

