

ANALIZA WRAŻLIWOŚCI MODELU POTĘGOWEGO LICZBY ZDARZEŃ DROGOWYCH W POLSCE¹

ANDRZEJ ROGOWSKI

dr inż., Uniwersytet Technologiczno-
-Humanistyczny w Radomiu,
Wydział Transportu i Elektrotechniki,
26-600 Radom, ul. Malczewskiego 29,
tel. 48 361 77 85,
e-mail a.rogowski@uthrad.pl

Streszczenie. Wypadki drogowe stanowią istotny problem społeczny i ekonomiczny, stąd różnorodne działania zmierzające zarówno do zmniejszenia ich liczby, jak i ich skutków. Jednym z pierwszych modeli szacujących liczbę ofiar śmiertelnych w ruchu drogowym był model Smeeda będący modelem potęgowym o zmiennych objaśniających – liczbie pojazdów i liczbie ludności. Choć klasyczny model Smeeda, przynajmniej w warunkach polskich, jest nieprzydatny, to z parametrami estymowanymi metodą najmniejszych kwadratów jako model wykorzystujący klasyczne podejście do prognozy wypadków na podstawie zmiennych opisujących wystawienie na ryzyko, dość dobrze opisuje liczbę ofiar śmiertelnych i rannych w wypadkach drogowych, jak i liczbę wypadków drogowych. Jednak jego przydatność do celów prognostycznych warunkowana jest zmiennością w czasie parametrów modelu i wrażliwością modelu na zamianę wartości zmiennych objaśniających i parametrów modelu. W artykule przedstawiono wyniki analizy wrażliwości modeli potęgowych dla których parametry estymowano na podstawie danych rzeczywistych z lat 2000–2011. Stwierdzono małą wrażliwość modeli na zmiany wartości zmiennych objaśniających (uwzględniając zakres danych właściwy dla warunków polskich), co można uznać za zaletę w przypadku prognozowania – minimalizacja błędów związanych z prognozowaniem wartości zmiennych objaśniających. Jednocześnie stwierdzono bardzo dużą wrażliwość modeli na zmiany wartości parametrów – wykładników potęg, co powoduje duże trudności w doborze danych rzeczywistych do szacowania parametrów modeli. Zaproponowano modele z ustaloną wartością parametru jednej ze zmiennych objaśniających i parametrem w funkcji czasu dla drugiej zmiennej, podając postacie tych funkcji.

Słowa kluczowe: model potęgowy, bezpieczeństwo ruchu drogowego, prognozowanie

Wprowadzenie

Wypadki drogowe stanowią istotny problem społeczny i ekonomiczny, stąd różnorodne działania zmierzające zarówno do zmniejszenia ich liczby, jak i ich skutków. Jedną z metod, przynajmniej dla oceny skuteczności podejmowanych działań, jest prognozowanie liczby wypadków drogowych i liczby ich ofiar. Do pierwszych ogólnie znanych modeli szacujących liczbę ofiar śmiertelnych w ruchu drogowym należy model Smeeda uzależniający liczbę F ofiar śmiertelnych ruchu drogowego w danym roku na danym obszarze od liczby ludności i liczby zarejestrowanych pojazdów. Model Smeeda jest modelem potęgowym o dwu zmiennych objaśniających (V – liczba pojazdów, P – liczba ludności) postaci:

$$F = aV^bP^c \quad (1)$$

przy czym parametry są równe: $a = 0,0003$, $b = 1/3$, $c = 2/3$ ($F = 0,0003\sqrt[3]{V \cdot P^2}$).

Model ten² wzbudzał i nadal wzbudza szereg kontrowersji, choćby z tego powodu, że parametry modelu są stałe, niezależne od obszaru geograficznego. Tym niemniej nadal jest on, wraz z modyfikacjami – polegającymi głównie na estymacji parametrów modelu w odniesieniu do wyróżnionego obszaru³ analizowany i sprawdzana jest jego przydatność w różnych krajach. Należy tutaj wymienić prace: Adamsa [1], który analizował klasyczny model Smeeda i jego modyfikacje dla 20 krajów z całego świata, Korena [2] (m.in. model Smeeda dla 139 krajów dla roku 2007), Akgüngöra [3] dla miast w Turcji, Valli [4] i Chakraborty'ego [5] dla obszaru Indii – ten ostatni rozpatrywał również modele liczby wypadków i liczby rannych. Dość szeroką dyskusję na temat samego modelu, kontrowersji z nim związanych, jego modyfikacji oraz literaturę znajdzie czytelnik w pracy [6]. Bardziej rozbudowane modele liczby zdarzeń drogowych znaleźć można w pracach [7] i [8].

W warunkach polskich, przynajmniej dla okresu do roku 1998, klasyczny model Smeeda jest nieprzydatny. Z jego równania wynika bowiem, że wraz ze wzrostem liczby pojazdów, jak i liczby ludności, liczba ofiar śmiertelnych rośnie, np. w roku 2011 wg równania Smeeda liczba ofiar śmiertelnych w Polsce wyniosłaby 9899 [9]. Tymczasem, przynajmniej od 1997 roku, obserwujemy wyraźną tendencję spadku zarówno liczby ofiar śmiertelnych, jak i liczby wypadków i rannych przy bardzo szybkim wzroście liczby pojazdów i jednocześnie w miarę stabilnym poziomie liczby ludności. Powstaje pytanie, czy dla innych wartości parametrów model potęgowy (1) będzie dawał dostatecznie dokładne wyniki i czy model potęgowy może być stosowany również dla prognozowania liczby rannych w wypadkach, liczby wypadków i liczby kolizji jako model wykorzystujący klasyczne podejście do prognozy na podstawie zmiennych opisujących wystawienie na ryzyko. Założono, tak jak w klasycznym modelu Smeeda, że nie będą uwzględniane inne czynniki mające wpływ (potencjalnie) na wypadkowość i ofiarochłonność ruchu drogowego – rozpatrywana będzie jedynie liczba pojazdów silnikowych i liczba ludności. Interesuje nas bardzo prosty model o możliwie małej

² Równanie $F = 0,0003\sqrt[3]{V \cdot P^2}$ nazywać będziemy dalej klasycznym równaniem Smeeda.

³ Model (1), dla którego dokonuje się estymacji parametrów, nazywany bywa w literaturze modelem (modyfikacją) Andreassena.

liczbie zmiennych objaśniających, łatwych do pozyskania i prognozowania⁴. W tym celu dokonano estymacji parametrów modelu na podstawie szeregów czasowych różnej długości, analizowano również modele dla poszczególnych województw. Szczegółowe analizy przedstawiono w [9].

Analizując modele dla liczby wypadków i liczby rannych w wypadkach drogowych, zwrócono uwagę na dużą zbieżność parametrów modeli dla tych samych szeregów czasowych – przy jednak istotnych różnicach w liczbie wypadków i liczbie rannych oraz dużym zróżnicowaniu wartości parametrów modeli (dla wszystkich badanych wielkości) dla poszczególnych województw. Różnic należało się spodziewać, gdyż uwzględniano tylko dwie wielkości, nie analizując innych czynników, np. mobilności, zróżnicowania infrastruktury drogowej, uwarunkowań społeczno-gospodarczych. Tylko częściowo (poprzez zmianę długości szeregu czasowego) uwzględniono wpływ czasu na wartości parametrów modeli. Odpowiedź na pytanie, czy istnieje w miarę prosta zależność funkcyjna wartości parametrów modeli od czasu, może być istotna dla zwiększenia użyteczności modeli potęgowych dla prognozowania wskaźników bezpieczeństwa w ruchu drogowym. Dla prawidłowej odpowiedzi na to pytanie niezbędne jest ustalenie wrażliwości modeli potęgowych na zamiany wartości parametrów modeli – choćby do oceny dokładności szacowania parametrów (zbyt mała dokładność skutkuje nieprzydatnością oszacowań) – i zmiany wartości zmiennych objaśniających modeli (dla uzyskania wartości modeli dla okresów przyszłych musimy dokonać prognoz wartości zmiennych objaśniających). Zwróćmy uwagę, że mówiąc o „dużej zbieżności parametrów modeli”, ocenialiśmy zgodność parametrów „z punktu widzenia dnia codziennego”, gdzie różnica 0,001 dla liczb z przedziału (-1,2) jest różnicą zwykle nieistotną. W obszarze nauki taka ocena jest niewłaściwa. W przypadku modeli potęgowych, dla bardzo dużych wartości zmiennych (a liczba pojazdów i liczba ludności liczone są w dziesiątkach milionów), różnica w oszacowaniu parametrów modelu rzędu 0,001 może być bardzo istotna dla wartości zmiennych objaśnianych, a „istotność” zależy od celów badań (i przyjętych kryteriów). Oceny istotności można dokonać, analizując różnicę w wartościach zmiennych objaśnianych.

Badania i analiza wyników

W analizie wykorzystano modele potęgowe typu (1) dla liczby ofiar śmiertelnych, rannych i liczby wypadków w Polsce, szacując parametry modeli metodą największej wiarygodności w oparciu o dane rzeczywiste z lat 2000–2011⁵. Dane dla lat 2000–2012 pozyskano z [10].

⁴ W [9] rozpatrywano również modele potęgowe i liniowe z jedną zmienną objaśniającą – liczbą pojazdów silnikowych – uzyskując bardzo podobne wyniki jak dla modeli z dwiema zmiennymi objaśniającymi.

⁵ Dla tych lat autor miał dostęp do wszystkich danych, w tym dla poszczególnych województw. Wstępne dane o liczbie wypadków, ofiar rannych i śmiertelnych w roku 2012 pochodzą ze strony KG Policji. We wcześniejszych badaniach [9] wykazano, że modele potęgowe z liczbą pojazdów silnikowych i/lub liczbą ludności jako zmiennymi objaśniającymi są nieprzydatne do prognozowania liczby kolizji w ruchu drogowym, stąd nie analizowano takiego modelu.

Uzyskano następujące modele⁶:

$$L_s = 1,0000 * V^{-0,6657} * P^{1,1278} \quad (2)$$

$$L_r = 1,0000 * V^{-0,5864} * P^{1,1925} \quad (3)$$

$$L_w = 1,0000 * V^{-0,5812} * P^{1,1743} \quad (4)$$

gdzie:

V – liczba zarejestrowanych pojazdów silnikowych w Polsce na koniec danego roku,

P – liczba ludności Polski na koniec danego roku,

L_s – liczba ofiar śmiertelnych wypadków drogowych w danym roku w Polsce,

L_r – liczba rannych w wypadkach drogowych w danym roku w Polsce,

L_w – liczba wypadków drogowych w danym roku w Polsce.

Obliczone dla modeli wartości współczynnika R^2 wynoszą odpowiednio: 0,704, 0,786, 0,791.

Analizowano zmianę wartości zmiennych objaśnianych przy zmianach wartości parametrów wg następującego schematu:

- dla wybranych wartości parametru b i c (oddzielnie), przy ustalonych pozostałych parametrach i rzeczywistych wartościach zmiennych objaśnianych dla lat 2000–2011 i wartościach prognozowanych dla roku 2012; skok zmian wartości parametru przyjęto 0,001, obszar zmienności przyjęto tak, by estymowana wartość parametru modeli była środkiem (w przybliżeniu) obszaru zmienności;
- dla wybranych kombinacji wartości parametrów b i c , przy ustalonym parametrze a i wartościach zmiennych objaśnianych jak w pkt. a; skok zmian wartości parametru przyjęto: $\pm 0,001$; $\pm 0,002$; $\pm 0,003$ w stosunku do wartości estymowanych parametrów modeli;
- zmiany wartości zmiennych prognozowanych dla wybranych kombinacji wartości zmiennych objaśnianych V i P (łącznie); skok zmian wartości (V,P) przyjęto: $(\pm 100000; \pm 5000)$, $(\pm 200000; \pm 10000)$, $(\pm 300000; \pm 15000)$.

Nie analizowano wpływu zmian wartości parametru a , gdyż jego wpływ jest oczywisty – zmiana o $x\%$ współczynnika a powoduje zmianę wartości zmiennej objaśnianej o $x\%$. W przypadku modeli (2), (3), (4) (i wszystkich rozpatrywanych w dalszej części) zmiany wartości zmiennych objaśnianych mogłyby być istotne dopiero przy zmianach wartości parametru a rzędu 10^{-3} , a estymowane wartości parametru różnią się od 1 o mniej niż 10^{-6} , stąd przyjęto $a = 1$.

W dalszej części analizowano wartości parametrów modeli w funkcji czasu.

⁶ Zwróćmy uwagę, że uzyskane oszacowania parametrów modeli zdecydowanie różnią się od wartości parametrów występujących w klasycznym równaniu Smeeda.

Liczba ofiar śmiertelnych

Zmiany wartości zmiennej objaśnianej – liczby ofiar śmiertelnych – w sposób oczywisty zależą od wartości zmiennych objaśniających. Wartości rozpatrywanego modelu w funkcji czasu (liczonego w latach) dla danych rzeczywistych tworzą ciąg ściśle malejący dla każdej z rozpatrywanych wartości parametru c . Wyniki obliczeń zawiera rysunek 1. Jeśli rozpatrywać wartości modelu jako wartości funkcji dwuwymiarowej $L_f(r,c)$, gdzie r – rok, c – parametr modelu, to funkcja ta jest ściśle malejąca po współrzędnej r i ściśle rosnąca po współrzędnej c , np. dla $r = 2000$ przyrost wartości L_f (wraz ze wzrostem c od 1,119 do 1,37 co 0,001; przyrost „prawie” liniowy) wynosi od 94 do 126 (średnia 109), dla $c = 1,119$ spadek wartości L_f (wraz ze wzrostem r) wynosi od 94 do 64 (średnia 78), dla $c = 1,137$ odpowiednio od 126 do 86 (średnia 105)⁷. Analogicznie zachowuje się mo-

del ze względu na parametr b (rys. 2). Funkcja $L_f(r,b)$ jest funkcją ściśle malejąca po współrzędnej r i ściśle rosnąca po współrzędnej b , np. dla $r = 2000$ przyrost wartości L_f (wraz ze wzrostem b od -0,675 do -0,656; przyrost „prawie” liniowy) wynosi od 88 do 117 (średnia 102), dla $b = -0,675$ spadek wartości L_f (wraz ze wzrostem r) wynosi od 88 do 62 (średnia 75), dla $b = -0,656$ odpowiednio od 117 do 83 (średnia 99).

Można więc uznać, że wrażliwość modelu jest jednako- wa ze względu na każdy z parametrów b, c , przy czym należy ją uznać za wysoką – zmiana wartości każdego z parametrów o 0,01 powoduje zmianę wartości L_f o około 1000. Fakt, że wrażliwość modelu jest prawie identyczna dla obydwu parametrów powoduje, że zmiany przeciwstawne parametrów znoszą zmiany wartości L_f , zmiany zgodne su-

Tabela 1

Wrażliwość modeli potęgowych L liczby ofiar śmiertelnych, rannych i wypadków na zmiany wartości parametrów b i c													
Rok	b	+0,001	+0,001	-0,001	-0,001	+0,002	+0,002	-0,002	-0,002	+0,003	+0,003	-0,003	-0,003
	c	+0,001	-0,001	+0,001	-0,001	+0,002	-0,002	+0,002	-0,002	+0,003	-0,003	+0,003	-0,003
liczba zabitych: a = 1,0000, b = -0,6657, c = 1,1278													
2000	6207	6421	6201	6213	6000	6643	6195	6219	5800	6872	6188	6226	5606
2001	6030	6239	6025	6036	5829	6454	6019	6042	5634	6677	6013	6048	5446
2002	5817	6019	5812	5822	5623	6227	5807	5828	5435	6442	5802	5833	5253
2003	5721	5919	5716	5726	5530	6124	5711	5731	5345	6336	5706	5736	5166
2004	5534	5726	5529	5539	5349	5924	5525	5543	5169	6130	5520	5548	4996
2005	5506	5697	5502	5511	5322	5895	5497	5515	5143	6099	5493	5520	4971
2006	5251	5433	5247	5255	5074	5622	5243	5258	4904	5817	5239	5262	4739
2007	4988	5162	4985	4991	4820	5341	4981	4995	4658	5528	4978	4998	4501
2008	4696	4860	4693	4699	4538	5030	4691	4702	4384	5206	4688	4704	4236
2009	4602	4763	4600	4605	4447	4930	4597	4607	4297	5102	4595	4610	4151
2010	4514	4673	4512	4517	4362	4836	4510	4519	4214	5006	4508	4521	4072
2011	4371	4525	4369	4373	4223	4683	4367	4375	4080	4847	4365	4377	3942
2012	4226	4375	4225	4228	4083	4528	4223	4230	3944	4688	4221	4232	3811
liczba rannych: a = 1,0000, b = -0,5864, c = 1,1925													
2000	70818	73261	70747	70889	68456	75789	70677	70959	66173	78404	70606	71030	63966
2001	69034	71419	68968	69100	66729	73887	68903	69166	64501	76439	68837	69232	62347
2002	66873	69187	66813	66933	64637	71581	66753	66994	62475	74058	66693	67054	60385
2003	65890	68171	65832	65948	63685	70532	65774	66005	61553	72974	65717	66063	59493
2004	63982	66201	63929	64035	61838	68496	63876	64088	59765	70871	63823	64141	57762
2005	63692	65901	63640	63744	61557	68187	63588	63796	59493	70551	63535	63849	57499
2006	61070	63193	61025	61116	59019	65389	60979	61162	57037	67662	60933	61208	55121
2007	58368	60401	58329	58407	56403	62505	58290	58447	54505	64683	58251	58486	52670
2008	55355	57288	55323	55387	53487	59289	55291	55419	51682	61360	55259	55452	49937
2009	54388	56290	54358	54418	52551	58258	54328	54448	50776	60295	54299	54478	49060
2010	53574	55450	53547	53602	51761	57392	53519	53629	50010	59402	53491	53657	48318
2011	52076	53903	52052	52101	50312	55793	52028	52125	48607	57750	52004	52149	46960
2012	50553	52329	50532	50574	48838	54167	50512	50595	47181	56069	50491	50616	45580
liczba wypadków: a = 1,0000, b = -0,5812, c = 1,1743													
2000	56078	58013	56022	56134	54207	60014	55966	56190	52400	62085	55910	56246	50652
2001	54678	56567	54626	54730	52852	58521	54574	54782	51087	60543	54522	54835	49381
2002	52981	54815	52934	53029	51209	56711	52886	53077	49497	58674	52838	53125	47841
2003	52209	54017	52164	52255	50462	55888	52118	52301	48773	57823	52072	52347	47141
2004	50711	52469	50669	50753	49011	54289	50627	50795	47369	56171	50585	50837	45781
2005	50483	52234	50442	50525	48791	54046	50401	50566	47156	55920	50359	50607	45575
2006	48424	50107	48387	48460	46797	51848	48351	48496	45226	53650	48315	48533	43707
2007	46300	47912	46269	46331	44741	49581	46238	46362	43235	51309	46206	46393	41780
2008	43930	45464	43904	43955	42447	47052	43879	43981	41015	48696	43853	44007	39630
2009	43169	44678	43145	43193	41711	46240	43122	43217	40302	47857	43098	43240	38940
2010	42525	44015	42504	42547	41087	45556	42482	42569	39697	47151	42460	42591	38353
2011	41347	42797	41328	41366	39946	44298	41308	41385	38593	45851	41289	41405	37285
2012	40148	41558	40132	40165	38786	43018	40115	40182	37470	44529	40098	40198	36199

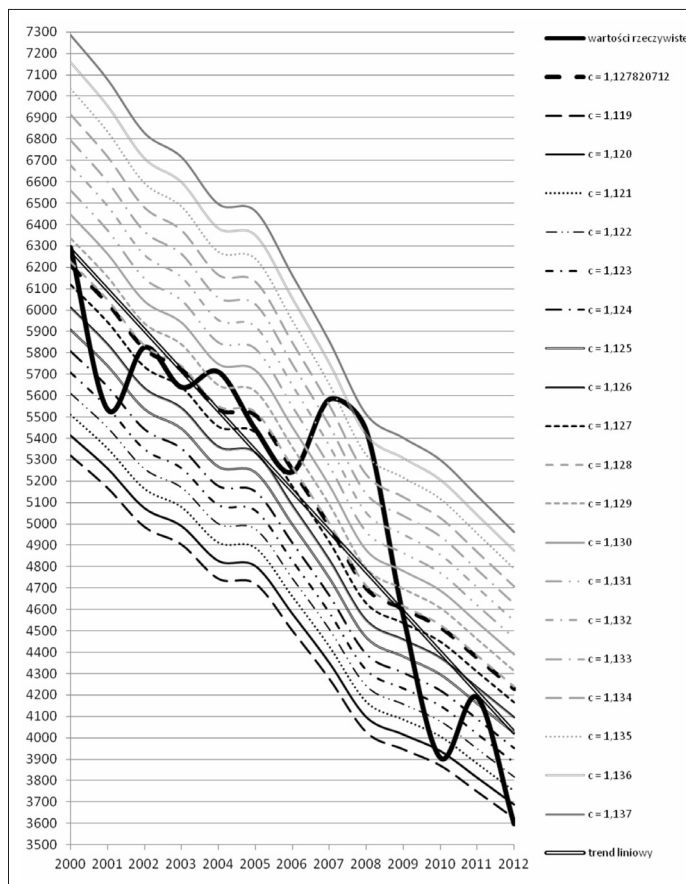
Źródło: opracowanie własne na podstawie [10]

⁷ Należy pamiętać, że zmiany liczby ludności w poszczególnych latach w okresie 2000–2011 nie były monotoniczne w przeciwieństwie do zmiany liczby pojazdów.

mują się. Można uznać, że są addytywne (przy zmianach nie większych niż 0,01). Poglądowe wyniki obliczeń zawiera tabela 1⁸, ilustrację wyników rys. 1 i 2.

Wrażliwość modelu na zmiany wartości zmiennych objaśniających można uznać za bardzo małą (nieistotną) – np. błąd szacunku (który należy uznać za duży) liczby pojazdów o 300 tys. i jednoczesny błąd liczby ludności o 15 tys. (w tym przypadku „błąd się sumuje”) skutkuje błędem zmiennej objaśnianej ok. 35, przy wartości zmiennej rzędu 4,2 tys. (tab. 2).

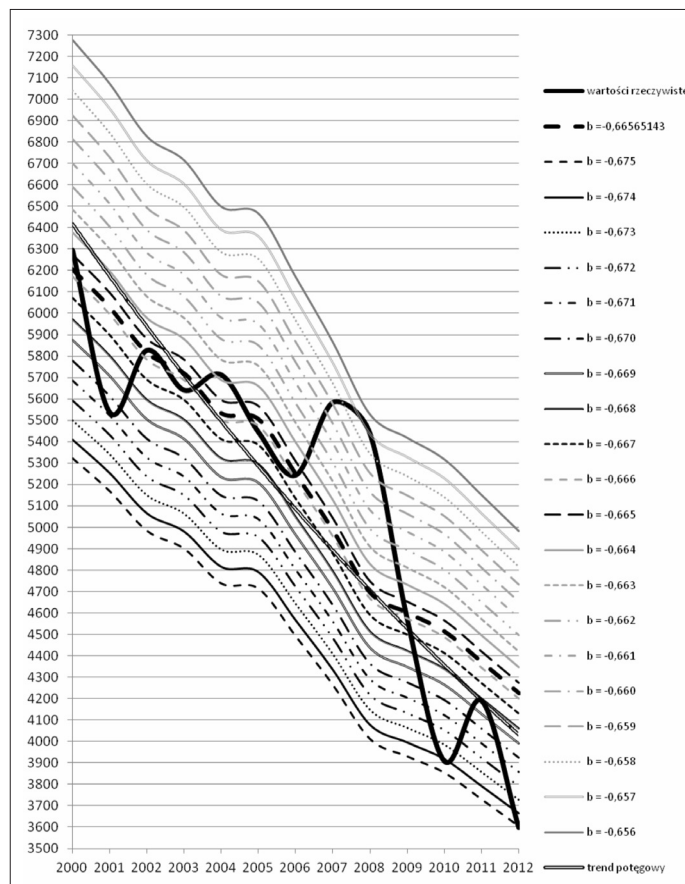
Analizując rysunki 1 i 2 wydaje się, że można tak dobrać zakres parametrów modelu, by rzeczywiste i prognozowane wartości L_s (analogicznie sytuacja przedstawia się dla L_r i L_w) zawierały się między wartościami zmiennej objaśnianej dla skrajnych wartości parametrów. Jednak użyteczność takiego oszacowania jest znikoma, gdyż różnica między oszacowaniem górnym a dolnym wynosi ok. 40% wartości zmiennej objaśnianej.



L_s – liczba ofiar śmiertelnych wypadków drogowych w danym roku w Polsce, V – liczba pojazdów silnikowych w danym roku w Polsce, P – liczba ludności w danym roku w Polsce; parametry modelu estymowano metodą najmniejszych kwadratów na podstawie danych z lat 2000–2011, otrzymując równanie: $L_s = 1,0000 * V^{-0.6657} * P^{1.1278}$; równanie trendu liniowego $y = -188,78x + 383843, R^2 = 0,799$

Rys. 1. Wrażliwość modelu potęgowego $L_s = aV^bP^c$ liczby ofiar śmiertelnych w ruchu drogowym w Polsce w latach 2000–2011 na zmiany parametru c

Źródło: opracowanie własne na podstawie [10]



L_s – liczba ofiar śmiertelnych wypadków drogowych w danym roku w Polsce, V – liczba pojazdów silnikowych w danym roku w Polsce, P – liczba ludności w danym roku w Polsce; parametry modelu estymowano metodą najmniejszych kwadratów na podstawie danych z lat 2000–2011, otrzymując równanie: $L_s = 1,0000 * V^{-0.6657} * P^{1.1278}$; równanie trendu potęgowego $y = 2,3 \cdot 10^{260} \cdot x^{-71,72}, R^2 = 0,775$

Rys. 2. Wrażliwość modelu potęgowego $L_s = aV^bP^c$ liczby ofiar śmiertelnych w ruchu drogowym w Polsce w latach 2000–2011 na zmiany parametru b

Źródło: opracowanie własne na podstawie [10]

Tabela 2

Wrażliwość modeli potęgowych $L = aV^bP^c$ liczby ofiar śmiertelnych, rannych i wypadków na zmiany wartości zmiennych objaśniających V i P													
Liczba pojazdów [tys.]	25450	25550	25550	25350	25350	25650	25650	25250	25250	25750	25750	25150	25150
Liczba ludności [tys.]	38542	38547	38537	38547	38537	38552	38532	38552	38532	38557	38527	38557	38527
Liczba zabitych	4226	4216	4215	4238	4237	4206	4203	4250	4247	4195	4192	4262	4258
Liczba rannych	50553	50445	50430	50678	50662	50338	50306	50804	50772	50231	50184	50930	50883
Liczba wypadków	40148	40063	40051	40246	40234	39978	39954	40345	40321	39894	39858	40444	40407

Tabela zawiera prognozowane liczby ofiar śmiertelnych, rannych i liczby wypadków w roku 2012 w zależności od przyjętych (prognozowanych) wartości zmiennych objaśniających – liczby ludności i liczby pojazdów silnikowych.

Oznaczenia i wartości parametrów modeli jak dla równań (2), (3), (4)

Źródło: opracowanie własne

⁸ Istotne wnioski (również dla innych rozpatrywanych wielkości) można wyciągnąć z analizy wykresów przedstawionych na rysunku 3, omówienie którego znajduje się w dalszej części artykułu.

Liczba rannych i liczba wypadków

Modele potęgowe liczby rannych w wypadkach drogowych i liczby wypadków wykorzystują te same dane liczbowe, jednak wartości zmiennych objaśniających są o rząd większe niż w przypadku liczby ofiar śmiertelnych. Tymczasem wartości parametrów modelu wzrosły o około: parametr b o 11,9% i parametr c o 5,7% w przypadku liczby rannych i odpowiednio o 12,7% i 4,1% w przypadku liczby wypadków. Natomiast parametry modelu L_w w stosunku do modelu L_r zmalały o odpowiednio 0,9% i 1,5%, co skutkowało spadkiem wartości zmiennych objaśniających o ok. 20%. Jest to kolejne potwierdzenie dużej wrażliwości modelu potęgowego na zmianę parametrów modelu. Przypomnijmy, że parametr a praktycznie nie ma wpływu na wartości zmiennych objaśnianych (dla wartości przyjmowanych dla tych modeli).

Rozpatrując wartości modeli L_r i L_w , analogicznie jak modelu L_p , jako wartości funkcji dwuwymiarowych czasu r i parametrów odpowiednio b lub c , obserwujemy taki sam charakter zmian z uwzględnieniem faktu, że przyrosty są o rząd większe niż w przypadku modelu L_p , np.:

a) dla liczby rannych:

- dla parametru b : dla $r = 2000$ przyrost wartości L_r (wraz ze wzrostem b od -0,595 do -0,576 co 0,001; przyrost „prawie” liniowy) wynosi od 1020 do 1372 (średnia 1188), dla $r = 2012$ od 751 do 1020 (średnia 879); dla $b = -0,595$ spadek wartości L_r (wraz ze wzrostem r do 2012) wynosi od 1020 do 751 (średnia 883), dla $b = -0,576$ odpowiednio od 1372 do 1020 (średnia 1193),
- dla parametru c : dla $r = 2000$ przyrost wartości L_r (wraz ze wzrostem c od 1,184 do 1,202 co 0,001; przyrost „prawie” liniowy) wynosi od 1075 do 1446 (średnia 1252), dla $r = 2012$ od 768 do 1033 (średnia 894); dla $c = 1,184$ spadek wartości L_r (wraz ze wzrostem r do 2012) wynosi od 768 do 1033 (średnia 917), dla $c = 1,202$ odpowiednio od 1446 do 1033 (średnia 1234),

b) dla liczby wypadków:

- dla parametru b : dla $r = 2000$ przyrost wartości L_w (wraz ze wzrostem b od -0,590 do -0,571 co 0,001; przyrost „prawie” liniowy) wynosi od 805 do 1066 (średnia 930), dla $r = 2012$ od 594 do 794 (średnia 690); dla $b = -0,590$ spadek wartości L_w (wraz ze wzrostem r do 2012) wynosi od 805 do 594 (średnia 698), dla $b = -0,571$ odpowiednio od 1066 do 794 (średnia 927),
- dla parametru c : dla $r = 2000$ przyrost wartości L_w (wraz ze wzrostem c od 1,164 do 1,182 co 0,001; przyrost „prawie” liniowy) wynosi od 826 do 1111 (średnia 962), dla $r = 2012$ od 591 do 796 (średnia 689); dla $c = 1,164$ spadek wartości L_w (wraz ze wzrostem r do 2012) wynosi od 826 do 591 (średnia 705), dla $c = 1,182$ odpowiednio od 1111 do 796 (średnia 949),

Tak jak w przypadku modelu L_p można uznać, że wrażliwość modeli L_r i L_w jest jednakowa ze względu na każdy z pa-

rametrów b , c , przy czym należy ją uznać za wysoką – zmiana wartości każdego z parametrów o 0,01 powoduje zmianę wartości o około 10 000 dla L_r i ok. 8 000 dla L_w – i przyjąc, że zmiany przeciwstawne parametrów znoszą zmiany wartości modeli, zmiany zgodne sumują się. Można uznać, że są addytywne. Poglądowe wyniki obliczeń zawiera tabela 1.

Wrażliwość modeli na zmiany wartości zmiennych objaśniających można uznać za bardzo małą (nieistotną) – błąd szacunku liczby pojazdów o 300 tys. i jednoczesny błąd liczby ludności o 15 tys. (w tym przypadku „błąd się sumuje”) skutkuje błędem ok. 400 dla L_r , przy wartości zmiennej objaśnianej rzędu 55,5 tys. i ok. 300 dla L_w , przy wartości zmiennej objaśnianej rzędu 40 tys. (tab. 2).

Ponieważ przebieg wykresów wartości rzeczywistych uzyskanych z modeli potęgowych (uwzględniając, że wartości są około 10 razy większe) mają taki sam charakter jak dla liczby ofiar śmiertelnych (w latach 2000–2006 wahania są mniejsze), nie zamieszczano wykresów dla tych wielkości.

Uogólnienie rozważań

Rozważania dotyczące liczby ofiar śmiertelnych, rannych i wypadków można uogólnić, analizując błąd względny. Rozważmy model (1) i założmy wzrost jednego z parametrów b , c o wartość x , np. b , przy niezmiennych wartościach pozostałych parametrów. Jeśli (w %) zmieni się wartość F , mamy:

$$\Delta F = \frac{(aV^{b+x}P^c - aV^bP^c)}{aV^bP^c} \cdot 100 = \frac{aV^bP^c(V^x - 1)}{aV^bP^c} \cdot 100 = (V^x - 1) \cdot 100 \quad (5)$$

Ponieważ założono małe zmiany wartości parametru (x bliskie 0), więc V^x jest bliskie 1 (większe od 1 dla $x > 0$ i mniejsze od 1 dla $x < 0$). Analogiczny wynik uzyskamy dla parametru c i zmiennej P . Przyrost względny ΔF zależy więc (dla ustalonej wartości x) od wartości zmiennej objaśniającej, jednak dla małych zmian x , nawet przy dużych zmianach wartości V (lub P), różnice w przyrostach są małe – rzędu 0,05 pkt. procentowego dla $x = 0,001$, natomiast sam przyrost jest rzędu 1,75%. Przykładowe obliczenia przedstawiono na rysunku 3 (krzywe 4, 5, 6 dla zmiennej V i 7, 8, 9 dla zmiennej P).

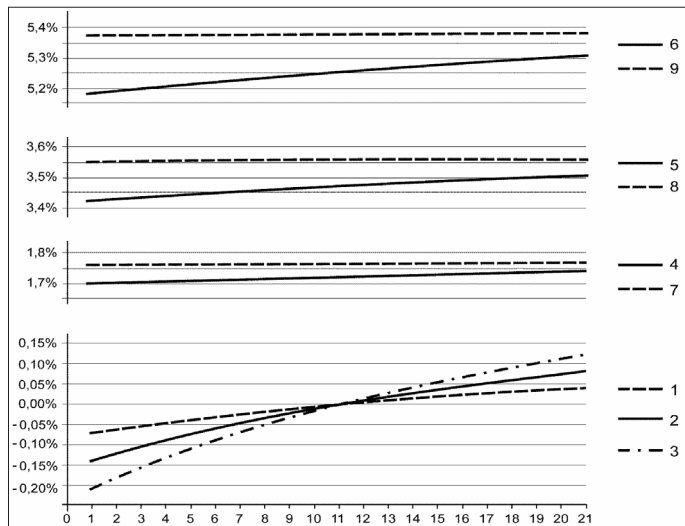
W przypadku jednoczesnej zmiany wartości parametrów b o $-x$ i c o x otrzymujemy:

$$\Delta F = \frac{(aV^{b-x}P^{c+x} - aV^bP^c)}{aV^bP^c} \cdot 100 = \frac{aV^bP^c(V^{-x}P^x - 1)}{aV^bP^c} \cdot 100 = \left(\left(\frac{P}{V} \right)^x - 1 \right) \cdot 100 \quad (6)$$

Przyrost zależy więc od stosunku zmiennych objaśniających (P/V dla $x > 0$ i V/P dla $x < 0$)⁹. Przykładowe obliczenia przedstawiono na rysunku 3 (krzywe 1, 2, 3). Z obliczeń wynika, że dla $x = 0,003$ zmiana wynosi od ok. -0,21% do 0,12% dla stosunku zmiennych objaśniających od 0,5 do 1,5.

⁹ Zmieniając parametry b i c , jednocześnie o x uzyskujemy $((VP)^x - 1) \cdot 100$ i dla danych jak na rysunku 3 zmianę rzędu 3,5% dla zmiany x o 0,001 (więc np. dla $x = 0,003$ rzędu 11%).

Istotnym wnioskiem jest, uwzględniając wykazaną wcześniej „addytywność” zmian wartości zmiennej objaśnianej ze względu na zmiany parametrów b i c , że parametry b i c rozpatrywanych modeli potęgowych należy obliczać z dokładnością do 4 miejsc po przecinku. Zaokrąglenie do 3 miejsc po przecinku może skutkować błędem rzędu 1,7% wartości zmiennej objaśnianej¹⁰.



Nr =	Nr*																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
P/V	0,5	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9	0,95	1	1,05	1,1	1,15	1,2	1,25	1,3	1,35	1,4	1,45	1,5
V	2050	2100	2150	2200	2250	2300	2350	2400	2450	2500	2550	2600	2650	2700	2750	2800	2850	2900	2950	3000	3050
P	3815	3818	3821	3824	3827	3830	3833	3836	3839	3842	3845	3848	3851	3854	3857	3860	3863	3866	3869	3872	3875

* Nr – kategoria umowna, oznacza kolejną wartość zmiennej (P/V) dla funkcji 1, 2, 3; $V \cdot 10^4$ dla funkcji 4, 5, 6 (np. Nr = 1 oznacza 20500000); $P \cdot 10^4$ dla funkcji 7, 8, 9 (np. Nr = 1 oznacza 38150000);
 1, 2, 3 – zmiana wartości modelu potęgowego ΔF w funkcji P/V (wzór (6)) dla przyrostu x odpowiednio: 0,001, 0,002, 0,003;
 4, 5, 6 – zmiana wartości modelu potęgowego ΔF w funkcji V (wzór (5)) dla przyrostu x odpowiednio: 0,001, 0,002, 0,003;
 7, 8, 9 – zmiana wartości modelu potęgowego ΔF w funkcji P (wzór (5)) dla przyrostu x odpowiednio: 0,001, 0,002, 0,003;

Rys. 3. Względna zmiana wartości modeli potęgowych (1) w zależności od wartości zmiennej objaśniającej i zmiany wartości parametrów
 Źródło: opracowanie własne

Czynnik czasu

Najprostszym sposobem uwzględnienia czynnika czasu jest wyznaczenie trendu. Wyznaczono trendy liniowy ($y = ax + b$) i potęgowy ($y = ax^b$) na podstawie danych rzeczywistych z lat 2000–2011, uzyskując następujące równania trendu i współczynniki R^2 (dla liczby ofiar śmiertelnych linie trendu zaznaczono na rys. 1 – trend liniowy, rys. 2 – trend potęgowy)¹¹:

- dla liczby ofiar śmiertelnych

$$y = -188,78x + 383843, R^2 = 0,799; \quad y = 2,3 \cdot 10^{260} \cdot x^{-77,72}, R^2 = 0,775$$

- dla liczby ofiar rannych

$$y = -1919,68x + 3911020, R^2 = 0,880; \quad y = 2,3 \cdot 10^{224} \cdot x^{-66,5}, R^2 = 0,858$$

- dla liczby wypadków

$$y = -1497,41x + 3051532, R^2 = 0,895; \quad y = 6,8 \cdot 10^{218} \cdot x^{-64,85}, R^2 = 0,877$$

¹⁰ Jeżeli z jakiegos powodu błąd rzędu 1,7% nie jest istotny, można podawać z mniejszą dokładnością.

¹¹ Wszystkie obliczenia wykonano w Excelu.

gdzie:

- x – czas – rok, dla którego dokonuje się obliczenia, $x = 1998, 1999, \dots$,
- y – liczba odpowiednio ofiar śmiertelnych, rannych i wypadków w roku, dla którego dokonuje się obliczenia.

Model trendu liniowego zaskakująco dobrze dopasowany jest do liczby wypadków (dla lat 1998–2012 tylko 4 razy różnica między wartością rzeczywistą a wartością modelu przekraczała 5% – jednak mniej niż 9%) i liczby rannych (odpowiednio 5 razy, mniej niż 9,5%). Podobnie sytuacja wygląda w przypadku trendu potęgowego, z tym jednak, że dwukrotnie błąd przekroczył 10% (zawyżone dla liczby wypadków dla roku 1999 – 11,2% i dla 2010 – 10,6%, dla liczby rannych odpowiednio: -12,6% dla roku 2007 i -13,9% dla roku 2008). W przypadku liczby zabitych dopasowanie obu trendów jest znacznie gorsze (pięciokrotnie różnica przekraczała 10%) i wydaje się mało przydatne. Wynika to m.in. z większej zmienności liczby ofiar śmiertelnych wypadków w funkcji czasu.

Innym podejściem do uwzględnienia czasu jest estymowanie parametrów modeli z uwzględnieniem szeregów czasowych o różnej długości i dokonywania prognoz krocących. Wykorzystano szeregi czasowe długości 4, 6 i 8 i dokonano estymacji parametrów modeli i prognoz dla lat 2006–2012¹². Przykładowe wyniki dla liczby rannych w wypadkach zawiera tabela 3 (w przypadku liczby ofiar śmiertelnych zróżnicowanie wartości parametrów jest jeszcze większe). Podejście takie okazało się nieprzydatne zarówno dla celów prognozowania – w niektórych przypadkach błąd przekraczał 25%, jak dla próby uchwycenia zależności parametrów od czasu¹³.

Tabela 3

Parametry b i c modelu $L_r = aV^bP^c$ estymowane w oparciu o szeregi czasowe długości 4, 6, 8 i prognozy liczby rannych w wypadkach drogowych w Polsce w latach 2006–2012										
Rok	Dane rzeczywiste	I			II			III		
		b	c	prognoza	b	c	prognoza	b	c	prognoza
2006	59123	-0,7901	1,3855	58997	-0,7247	1,3232	59327	-0,5821	1,1877	60293
2007	63224	-0,6435	1,2456	56559	-0,7053	1,3047	56212	-0,5325	1,1403	57927
2008	62097	-0,0470	0,6771	61467	-0,2946	0,9142	59104	-0,4618	1,0740	57210
2009	56046	0,1414	0,4959	62788	-0,0662	0,6957	61520	-0,2410	0,8636	59983
2010	48952	-0,1743	0,7983	59157	-0,2246	0,8468	58873	-0,2992	0,9186	58303
2011	49501	-1,5070	2,0846	48509	-0,4925	1,1027	53526	-0,4830	1,0936	53627
2012	45792	-2,1005	2,6610	43554	-0,8122	1,4112	49423	-0,6181	1,2232	50629

W tabeli podano wartości parametrów b i c (parametr $a = 1$) modelu potęgowego liczby rannych w wypadkach drogowych w Polsce estymowane w oparciu o szeregi długości 4 – kolumna I, 6 – kolumna II i 8 – kolumna III oraz prognozy uzyskane w oparciu o te modele. Parametry modelu dla roku r estymowano na podstawie danych rzeczywistych dla odpowiednio 4, 6 i 8 lat poprzedzających
 Źródło: opracowanie własne na podstawie danych [10]

¹² Od roku 1998 nastąpiła trwała zmiana tendencji – spadek, stąd racjonalnym jest wykorzystanie danych nie starszych niż z 1998 roku. Dla szeregu czasowego długości 8 dla roku 2006 można po raz pierwszy dokonać estymacji parametrów i prognozy.

¹³ Przynajmniej autor nie był w stanie takiej przydatnej zależności wychwycić.

Innym sposobem jest ustalenie wartości parametru jednej ze zmiennych objaśniających jako równe jeden i wyznaczenie parametru dla drugiej zmiennej w funkcji czasu. Otrzymamy wtedy dla każdej badanej zmiennej objaśniającej dwa typy modeli potęgowych:

$$L_z^1(x) = V(x)^{b_z(x)} * P(x) \quad (7)$$

$$L_z^2(x) = V(x) * P(x)^{c_z(x)} \quad (8)$$

gdzie:

x – rok, dla którego wykonuje się obliczenia ($x \geq 1998$),

$L_z^1(x), L_z^2(x)$; – liczba ofiar śmiertelnych ($z = s$), rannych ($z = r$) i wypadków ($z = w$) w roku x ,

$V(x)$ – liczba pojazdów motorowych na koniec roku x ,

$P(x)$ – liczba ludności na koniec roku x ,

$b_z(x), c_z(x)$ – parametry w funkcji czasu.

Dla poszczególnych modeli zaproponowano funkcje $b(x)$ i $c(x)$ – ich zestawienie oraz miary dopasowania R^2 zawiera tabela 4.

Tabela 4

Parametry b i c modeli (7), (8) w funkcji czasu		
Model	$b_z(x)$	R^2
$L_s^1(x)$	$b_s(x) = \begin{cases} -0,5195 - 0,0021(x - 1998), 1998 \leq x < 2004 \\ -0,5264 - 0,0021(x - 2004), x \geq 2004 \end{cases}$	0,884
$L_r^1(x)$	$b_r(x) = \begin{cases} -0,3737 - 0,0023(x - 1998), 1998 \leq x < 2006 \\ -0,3824 - 0,0023(x - 2007), x \geq 2007 \end{cases}$	0,859
$L_w^1(x)$	$b_w(x) = \begin{cases} -0,3869 - 0,0023(x - 1998), 1998 \leq x < 2006 \\ -0,3963 - 0,0023(x - 2007), x \geq 2007 \end{cases}$	0,888
	$c(x)$	
$L_s^2(x)$	$c_s(x) = \begin{cases} -0,4290 - 0,0055(x - 1998), 1998 \leq x < 2004 \\ -0,5264 - 0,0021(x - 2004), x \geq 2004 \end{cases}$	0,910
$L_r^2(x)$	$c_r(x) = \begin{cases} -0,2920 - 0,0050(x - 1998), 1998 \leq x < 2005 \\ -0,3190 - 0,0060(x - 2005), x \geq 2005 \end{cases}$	0,922
	$c_{r1}(x) = -0,2920 - 0,0050(x - 1998), x \geq 1998$	0,874
$L_w^2(x)$	$c_w(x) = \begin{cases} -0,3020 - 0,0050(x - 1998), 1998 \leq x < 2005 \\ -0,3358 - 0,0052(x - 2005), x \geq 2005 \end{cases}$	0,916
	$c_{w1}(x) = -0,3020 - 0,0050(x - 1998), x \geq 1998$	0,903

Oznaczenia jak dla wzorów (7) i (8)

Źródło: opracowanie własne

Najlepsze efekty uzyskano w przypadku modelowania liczby wypadków i liczby rannych, przy czym modele (8) wykazują lepsze dopasowanie niż modele (7). Modele te łączą w sobie cechy wyznaczonych wcześniej modeli trendu (przy zaproponowanych zależnościach funkcyjnych parametrów) – trend spadkowy, jak i modeli potęgowych (2), (3), (4) – zmienność wynikającą ze zmian wartości zmiennych objaśniających, „podążanie” za wartościami rzeczywistymi. Zaproponowane funkcje są funkcjami sklejanymi (z funkcji liniowych), jednak w przypadku liczby rannych i wypadków dla modeli (8) funkcje liniowe (oznaczone w tab. 4 jako $C_{r1}(x), C_{w1}(x)$) również wykazywały bardzo

dobrze dopasowanie do danych empirycznych. Należy jednak pamiętać, że nadal są to bardzo uproszczone modele i w wielu przypadkach korzystniejsze (prostsze) będzie wykorzystanie modeli trendu lub modeli potęgowego lub liniowego z jedną zmienną objaśniającą – liczbą pojazdów silnikowych.

Podsumowanie

Z przeprowadzonych obliczeń i analiz wynika, że modele potęgowe L_p, L_r i L_w postaci $L = aV^bP^c$ są mało wrażliwe na zmiany wartości zmiennych objaśniających i bardzo wrażliwe na zmiany parametrów modelu – wykładników potęg oraz długość i zakres szeregu czasowego wykorzystanego do estymacji parametrów. Spowodowane jest to bardzo uproszczoną postacią modelu tylko z dwoma zmiennymi objaśniającymi (nie uwzględniającymi np. zmian infrastruktury, zachowań kierowców w ruchu, mobilności społeczeństwa itp.), ale wynika również częściowo z samej procedury estymacji parametrów. Mała wrażliwość na zmiany wartości zmiennych objaśniających może być zaletą w przypadku prognozowania wartości zmiennych objaśnianych – mniej istotne stają się błędy prognozowania wartości zmiennych objaśniających, jednak duża wrażliwość na różnice oszacowań parametrów modeli wynikające z przyjęcia różnych zestawów danych (długości szeregu czasowego) jest poważnym problemem znacznie ograniczającym użyteczność tych modeli. Wprowadzenie modyfikacji polegającej na ustaleniu jednego z parametrów jako równego jeden i związaniu drugiego parametru prostą funkcją czasu (zasadniczo niezależną od długości szeregu czasowego) znacznie podnosi ich użyteczność.

Literatura

- Adams J.G.U., *Smeed's Law: some further thoughts*, Traffic Engineering and Control 28 (2) 1987.
- Koren C., Borsos A., *Is Smeed's Law Still Valid? A World-Wide Analysis of the Trends in Fatality Rate*, Journal of Society for Transportation and Traffic Studies (JSTS) Vol.1 2010.
- Akgüngör A.P., Dogan E., *An application of modified Smeed, adapted Andreassen and artificial neural network accident models to three metropolitan cities of Turkey*, Scientific Research and Essay Vol. 4 (9), 2009.
- Chakraborty S., Roy S.K., *Traffic accident characteristics of Kolkata*, Transport and Communications Bulletin for Asia and the Pacific No. 74, 2005.
- Valli P., *Road accident models for large metropolitan cities of India*, IATSS Research Vol. 29 No.1, 2005.
- Szymanek A., *Teoria i metodologia zarządzania ryzykiem w ruchu drogowym*. Politechnika Radomska, Radom 2012.
- Jamroz K., *Metoda zarządzania ryzykiem w inżynierii drogowej*, Wydawnictwa Politechniki Gdańskiej, Gdańsk 2011.
- Global status report on road safety 2013. Supporting a decade of action*, World Health Organization, ISBN 978 92 4 156456 4.
- Rogowski A., *Number of Motor Vehicles as an Explanatory Variable in the Prediction of Traffic Incidents in Poland*, The Archives of Transport, Volume 25 (2013) Issue 1 (w druku).
- http://www.stat.gov.pl/bdl/app/strona.html?p_name=indeks (GUS, Bank Danych Lokalnych), http://statystyka.policja.pl/portal/st/1302/76562/Wypadki_drogowe_raporty_roczne.html (KG Policji).