

ANALYSIS OF THE INFLUENCE OF FACTORS DETERMINING TRAFFIC ON UNDESIRABLE EVENTS

ANALIZA WPŁYWU CZYNNIKÓW WARUNKUJĄCYCH RUCH DROGOWY NA POWSTAWANIE ZDARZEŃ NIEPOŻĄDANYCH

Jacek Sipa¹, Marek Młyńczak²

(1,2) Wrocław University of Technology, Faculty of Mechanical Engineering
Politechnika Wroclawska Wydział Mechaniczny
50-371 Wrocław ul. Łukasiewicza 5

e-mails: jacek.sipa@pwr.wroc.pl, marek.mlynczak@pwr.wroc.pl

Abstract. The increasing participation of an automobile as a means of transport causes the increase of traffic flow. The insufficient development of road infrastructure, maladjustment to the increasing number of vehicles, and the problem that drivers do not adjust to conditions existing in the system contribute to the low level of traffic safety. The attention has been given to factors that characterize the vehicle, the user and the environment. The example of assessment of the factors' influence on the safety level has been made on the basis of the obtained nonlinear, multiple regression models.

Keywords: road transport system, risk factors, safety, multiple regression models

Streszczenie. Coraz większy udział samochodu jako środka transportu powoduje coraz większe natężenie ruchu drogowego. Niedostateczny rozwój infrastruktury drogowej oraz niedostosowanie się kierowców do warunków panujących w systemie przyczyniają się niskiego poziomu bezpieczeństwa na drogach. Zwrócono uwagę na rolę czynników charakteryzujących system transportu. Oceny wpływu poszczególnych wielkości na poziom bezpieczeństwa obiektu dokonano na podstawie otrzymanych modeli regresyjnych. W analizie wykorzystano modele liniowe i nieliniowe regresji wielokrotnej

Słowa kluczowe: system transportu drogowego, czynniki ryzyka, bezpieczeństwo, regresja wielokrotna

ANALYSIS OF THE INFLUENCE OF FACTORS DETERMINING TRAFFIC ON UNDESIRABLE EVENTS

1. Introduction

The increasing participation of an automobile as a means of transport causes the increase of traffic flow. That situation should entail the development of the whole road transport system. In fact, the insufficient development of road infrastructure, maladjustment to the increasing number of vehicles, and the problem that drivers do not adjust to conditions existing in the system contribute to the low level of traffic safety. Although road accidents are the result of human behavior and choices, they are not planned and chosen by drivers. An accidental event is a result of connection of certain factors determining road traffic at a particular moment. When the driver does not adjust to an existing situation on the road, his actions may lead to a road accident. Drivers take their decisions in a subjective way, often being under the influence of a sudden change on a road section, impossible to be predicted earlier in a particular place and time. Of course, it can not be said that the driver does not have the ability to affect his participation in traffic. He always chooses a certain margin of safety, but he can not escape the accidental risk. The single event is the result of a random effect and it is impossible to predict when and where the accident is going to happen. The accidental level of risk can be estimated, but it does not mean the certainty of accident occurrence in a given place and at a given moment. Disposing of an accidental data that happened on a well known road section, it can be possible to describe by means of mathematical relationships the influence of road characteristics on an average number of accidental events or accidental rate factors [1, 2]. Great data sets often reveal regularities of traffic process and its after-effects depending on factors that describe the system state. When someone knows these relationships, then he can characterize the influence of the factors' changes on the traffic safety level in a given, long-time period (e.g. estimation of expectation number of accidents per year). The whole accidental events number variation on a given road section consists of systematic variation, which depends on causal factors characterizing the accidental risk, and a purely random variation, which results from a random nature of accidental phenomenon. Having a notion of connections between the systematic variation and the causal factors, the

expectation number of events can be influenced and controlled. Of course, there will still be an unavoidable amount of random variation that can not be controlled. While describing the influence of individual causal factors on the road accident number, econometric methods can be very useful. Multiple regression is the most common analysis, but some conditions and assumptions have to be accepted [2, 3, 5].

1.1 Multiple regression models in the road accident prediction process

When all the factors that influence the road system state and take an important part in the occurrence of undesirable events are recognized, then general regularity responsible for the systematic part of the variation can be identified. The influence of the causal factors' variability on the safety level in a certain road section is always estimated with a given probability. Multiple regression models are the basic means to identify and determine the influence of recognized independent variables on the road safety level. The appropriate determination of connection between the dependent variable (the number of road accidents or safety level factors based on the number of accidents) and the group of independent variables (e.g. traffic flow, velocity, structure of users) require an assumption of some necessary conditions such as [2]:

- defining the set of independent variables $X(t)$,
- defining the nature of probability law that accident counts must follow (it is important on account of random variation), which is connected with assuming the probability distribution of the dependent variable,
- assuming the proper structure of the regression function $y=f(X(t))$,
- effective estimation of the function parameters b_i on the basis of assumed probability distribution of the dependent variable.

The most common model of the multiple regression is conventional linear regression, which can be defined as follows:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1)$$

where Y is the dependent variable, x_i represents independent variables, β_i represents estimated parameters, and ε is an error term. Model (1) have not been found as appropriate for modeling the number of accident events, because the model form is not restrained from predicting negative values and originally is used for variables that follow the normal probability law.

For several years two forms of the probability distribution have been used as the proper ones in the case of describing the randomness of accident counts: the Poisson distribution and the Pascal distribution (negative binomial distribution), that in principle is the modification of the Poisson distribution [2, 3, 4]. When assuming the randomness of the variable followed by Poisson distribution, the regression function is defined as follows:

$$\lambda = e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)} \quad (2)$$

where λ is an expected value of the dependent variable, and remaining parameters and variables are assumed likewise in the formula (1).

Assumed form of the regression model in accordance with formula (2) describes properly the relationship between the dependent variable λ and the group of independent variables x_i only in the case when the condition of probability distribution is satisfied. In the case of Poisson distribution the constraint of equal average and variation (described by λ) is the necessary condition.

When modeling the number of road accidents the overdispersion difficulty is a problem that often occurs. Then the condition for the same value of the average and the variance is not satisfied. The more appropriate model that makes the overdispersion possible is Pascal distribution, defined also as a negative binomial model. Pascal distribution is the modification of Poisson distribution. It assumes that the Poisson parameter λ is itself random, drawn from a Gamma distribution. Then the regression model is defined as follows:

$$\lambda = e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \xi)} \quad (3)$$

where φ is the Gamma distributed error term and all other variables are defined as in formula (2). For a such assumed model the relationship between the expected value λ and variance σ is given by equation:

$$\sigma = \lambda(1 + \theta \cdot \lambda) \quad (4)$$

where ν is the overdispersion parameter. If ν is not significantly different from 0 (significance should be checked by a statistical test), then the regression analysis based on the Poisson model is appropriate.

2. Road traffic accident database and multiple regression analysis

In order to carry out regression analysis determining the relationship between the average number of road accidents per year LZ and the set of independent variables characterizing road traffic, the section of A-4 motorway situated on the territory of south-western province of Poland has been chosen. The database of road accidents in this section has been created on the basis of information obtained from the Provincial Command of the State Fire Service in Wrocław (the capital of the province). Information about factors determining road traffic has been received from The General Directorate for National Roads and Motorways. The analyzed section has been partitioned into subsections between junctions (exits and approaches of motorway) and according to measurements of daily average traffic flow (SDR). Subsections obtained in that way are presented in table 1. The scope of the events being described in the database covers years 1999-2005. In that period, information about 1015 road accidents was recorded.

Table 1. Analyzed section of the motorway with the division into eleven subsections

<i>Subsection</i>	<i>length [km]</i>	<i>Description of the subsection</i>	<i>SDR [veh/24h]</i>	
			<i>2000</i>	<i>2005</i>
I	10,8	Krzywa-Budziwojów	9201	12820
II	11,1	Budziwojów-Czerwony Kościół	9486	13702
III	8,2	Czerwony Kościół-Legnica	10405	14700
IV	12,5	Legnica-Wądroże.	11601	15754
V	7,5	Wądroże-Budziszów	11601	17772
VI	13,3	Budziszów -Kostomłoty	12377	17878
VII	11,9	Kostomłoty-Kąty Wrocławskie	16977	21293
VIII	15,9	Kąty Wrocławskie -Węzeł Bielany	18468	24027
IX	11,2	Węzeł Bielany -Krajków	13456	22716
X	14,2	Krajków-Brzezimierz	12943	22453
XI	14,7	Brzezimierz-Granica województwa	12818	21767

2.1 Assumed regression model and results of the analysis

While modeling the expected number of road accidents, the regression model based on the assumption that randomness of the dependent variable follows Poisson probability distribution. Each obtained regression relationship was checked from the point of view of the overdispersion

concerning the dependent variable, the expected number of accidents in a particular year (LZ). For that purpose, ν parameter was estimated and its statistical signification was checked. If the parameter was not significantly different from 0, then the Poisson model was considered as adequate. Multiple correlation coefficients were also calculated in order to check the variation value explained by models. Three different methods were used to calculate three coefficients: the squared multiple correlation coefficient R^2 , the weighted multiple correlation coefficient R_w^2 and the Freeman-Tukey correlation coefficient R_{ft}^2 , based on the variance stabilizing transformation of a Poisson variable. The corrected multiple correlation coefficients were also calculated in order to check the value of systematic variation explained by models on account of Poisson distribution, assumed to describe the randomness of dependent variable. MD statistic (Model Deviation) was estimated in order to check significance of the obtained regression model. In the assumed model, variables were combined by the formula:

$$\lambda_i = Lod_i^{b_1} \cdot SDR_i^{b_2} \cdot e^{\sum b_0 + b_j \cdot x_{ji}}, j > 2 \quad (5)$$

where variables and parameters denote:

b – estimated coefficients of particular variables,

Lod – length of a subsection [km],

SDR – daily average traffic flow of a given subsection [veh/24h],

x_j – additional causal variables such as participation of heavy-duty vehicles Uc [%], the number of all junctions (all exits and approaches of motorway) ZW , the number of road junctions ZW_D , the number of road junctions with national roads ZW_k , communal roads ZW_g , local roads ZW_l , the number of exits and approaches to parking lots ZW_p , modernization of subsection PoM (values from 0 to 1).

Additionally, logarithmic transformations of two variables Lod and SDR were applied. That allowed the dependent variable to obtain 0 if the length of the subsection or the daily average traffic flow was 0.

Particular stages of the model's formation, by including successive causal variables are presented in tables 2 and 3. Marked models (8-12) are characterized by the good statistical significance of assumption that the randomness of a dependent variable follows Poisson distribution. That means the parameters b_i were estimated in a proper way. The ν parameter is not significantly different from 0, and there is a high value of variation explained by the model. Insignificant parameters are written in italics. Obtained regression models can be used in prediction of road accident numbers when future values of independent variables (factors describing

road traffic) are known. However, while making prognosis, the following conditions have to be taken into consideration:

- obtained regression model treats only the considered road section, because the role of particular independent variables (parameters) in the model was defined for this section [2],
- it is hard to determine the constancy influence of the obtained models in a long-term prognosis,
- it should be remembered that usually only some part of the systematic variation is explained, and there would still be an unavoidable amount of random variation (in Poisson model the variance is given by λ), which can not be explained [1]. The expected number of events is estimated, the most probable but not sure of occurring.

Table 2. Obtained regression models

Model	Independent variable										
	<i>Const.</i>	<i>Lod</i>	<i>SDR</i>	<i>Uc</i>	<i>ZW</i>	<i>ZWD</i>	<i>ZWk</i>	<i>ZWg</i>	<i>ZWI</i>	<i>ZWp</i>	<i>PoM</i>
1	-7,06	1,65	0,56	-	-	-	-	-	-	-	-
2	-7,23	1,55	0,89	-0,09	-	-	-	-	-	-	-
3	-9,22	1,49	0,82	-	0,01	-	-	-	-	-	-
4	-5,19	2,25	0,60	-0,11	-0,05	-	-	-	-	-	-
5	-6,75	1,68	0,83	-0,09	-	-0,02	-	-	-	-	-
6	-2,33	2,45	0,22	-0,09	-	-	0,05	-0,16	-0,13	-	-
7	-4,56	2,60	0,43	-0,10	-	-	-0,02	-0,08	-0,07	-0,09	-
8	-9,35	2,16	0,90	-0,05	-0,05	-	-	-	-	-	-0,77
9	-9,94	2,15	0,99	-0,06	-	-0,09	-	-	-	-	-0,86
10	-9,83	2,26	0,95	-0,06	-	-	-0,07	-0,10	-0,10	-	-0,85
11	-11,49	2,36	1,12	-0,06	-	-	-0,12	-0,04	-0,06	-0,06	-0,80
12	-9,91	1,59	1,06	-0,05	-	-	-	-	-	-	-0,65

Table 3. Statistical significance of marked models

Model	Multiple correlation coefficient						Model deviance		Y parameter (NBM)	
	R^2	Rp^2	Rw^2	Rwp^2	Rft^2	$Rftp^2$	MD	p	Y	p
8	0,72	0,88	0,69	0,87	0,69	0,87	261,26	10^{-7}	0,034	0,1936
9	0,74	0,90	0,70	0,88	0,70	0,90	267,63	10^{-7}	0,027	0,2617
10	0,76	0,92	0,72	0,91	0,72	0,92	269,76	10^{-7}	0,020	0,3763
11	0,78	0,95	0,75	0,94	0,75	0,96	269,83	10^{-7}	0,011	0,6067
12	0,70	0,85	0,67	0,84	0,67	0,85	248,61	10^{-7}	0,053	0,1054

3. Summary

It is hard to determine the influence of particular factors that determine the road transport system on the safety level in the object and this task demands to collect a proper past accident data aggregated in an appropriate long period of time. The choice of independent variables must be guided by theory and professional experience, rather than by the necessity of the best fitting of the model to the observed data of a dependent variable. The assumed form of the regression model together with the obtained relationships prove that the independent variables have been compiled correctly. Provided that the recommended conditions are satisfied, the obtained models can be used in assessing expected numbers of road accidents and their safety factors in the future periods.

Scientific work is financed from means for science in the years 2007-2008 as the research project no. 4564/B/T02/2007/33.

References

1. Baruya A.: *Speed-accident relationships on European roads*, Transport Research Laboratory, U.K., 1998.
2. Evans A. W.: *Estimating transport fatality risk from past accident data*, Accident Analysis & Prevention, 35, 2003.
3. Miaou S.P., Lum H.: *Modeling vehicle accidents and highway geometric design relationships*, Accident Analysis & Prevention, 25, 1993.
4. Otto W.: *Ubezpieczenia majątkowe*, WNT, Warszawa, 2004.
5. Washington S.P., Karlaftis M., Mannering F.: *Statistical and econometric methods for transportation data analysis*, CRC, Boca Raton, FL, 2003.

ANALIZA WPŁYWU CZYNNIKÓW WARUNKUJĄCYCH RUCH DROGOWY NA POWSTAWANIE ZDARZEŃ NIEPOŻĄDANYCH

1. Wstęp

Wzrastający udział samochodu jako środka transportu powoduje coraz większe natężenie ruchu drogowego, co powinno pociągać za sobą rozwój całego systemu komunikacji drogowej. W rzeczywistości, niedostateczny rozwój infrastruktury drogowej nieprzystosowanej do większej liczby pojazdów oraz niedostosowanie się kierowców do warunków panujących w systemie są głównymi przyczynami niskiego poziomu bezpieczeństwa na drogach. Chociaż wypadki drogowe są rezultatem ludzkiego zachowania nie są planowane przez kierowców. Do wypadku dochodzi w wyniku zaistnienia splotu pewnych czynników określających w danej chwili ruch drogowy. Jeżeli prowadzący pojazd nie dostosuje się zaistniałej sytuacji drogowej doprowadza do zdarzenia wypadkowego. Decyzje kierowców podejmowane są w sposób subiektywny, często pod wpływem nagłej zmiany na drodze, której nie można było wcześniej przewidzieć co do miejsca lub czasu. Oczywiście nie należy twierdzić, że kierowca nie ma wpływu na swoje uczestnictwo w ruchu. Zawsze wybiera pewien margines bezpieczeństwa, jednak ryzyko wypadku go nie ominie.

Pojedynczy wypadek jest zdarzeniem losowym, którego nie uda się przewidzieć. Określić można ryzyko jego powstania, ale nie będzie to oznaczało pewności jego zaistnienia w danym miejscu i danym czasie. Dysponując danymi o wypadkach na znanych odcinkach drogowych możliwe staje się ilościowe opisanie wpływu charakterystyki danego odcinka na średnią liczbę wypadków lub ich wskaźniki [1, 2]. Duże zbiory danych często ujawniają prawidłowości procesu ruchu i jego następstw w zależności od czynników określających stan obiektu. Znając te zależności określić można wpływ ich zmian na bezpieczeństwo ruchu w danym, dłuższym okresie (np. oszacować oczekiwaną liczbę zdarzeń w roku). Cała zmienność liczby zdarzeń wypadkowych na pewnym odcinku drogowym składa się ze zmienności systematycznej, zależnej od czynników przyczynowych określających ryzyko powstania zdarzenia niepożądanego oraz zmienności losowej wynikającej z natury zjawiska. Znając wpływ danych czynników na zmienność wariancji systematycznej, można

kontrolować oczekiwaną liczbę zdarzeń. Oczywiście zawsze pozostanie wariancja losowa, na którą już nie można wpłynąć.

Przy określaniu wpływu poszczególnych czynników na liczbę zdarzeń drogowych lub ich wskaźników bardzo pomocne mogą być metody ekonometryczne. Szeroko stosowany jest rachunek regresji wielokrotnej po przyjęciu pewnych uwarunkowań [2, 3, 5].

1.2 Modele regresyjne prognozowania liczby zdarzeń drogowych

Kiedy poznane są czynniki, które odgrywają ważną rolę w powstawaniu zdarzeń niepożądanych, można zidentyfikować ogólne prawidłowości procesu odpowiedzialne za człon zmian systematycznych. Z określonym prawdopodobieństwem szacuje się wpływ zmian tych czynników na bezpieczeństwo ruchu w danym miejscu sieci drogowej. Podstawowym narzędziem wykorzystywanym do identyfikacji i określenia wpływu poznanych zmiennych są wspomniane modele regresyjne. Prawidłowe określenie związku pomiędzy zmienną zależną (liczba zdarzeń drogowych lub wskaźniki bezpieczeństwa) a grupą zmiennych niezależnych (np. natężenie ruchu drogowego, prędkość, struktura uczestników) wymaga przyjęcia następujących warunków, takich jak [2]:

- określenie zbioru zmiennych niezależnych $X(t)$,
- poznanie natury losowości zjawiska występowania zdarzeń wypadkowych (odpowiedzialnej za wariancję losową), co związane jest z podaniem rozkładu zmiennej losowej,
- przyjęcie odpowiedniej postaci funkcji regresyjnej $y=f(X(t))$,
- efektywne wyznaczenie parametrów funkcji $y=f(X(t))$ na podstawie przyjętego rozkładu zmiennej zależnej.

Podstawowym modelem w regresji wielokrotnej jest model liniowy, który wyraża się wzorem:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1)$$

gdzie Y jest zmienna zależna, x_i to zmienne niezależne, β_i to estymowane parametry, a ε jest błędem estymacji. Wzór (1) nie znalazł jednak zastosowania w modelowaniu liczb zdarzeń drogowych, ponieważ z założenia jest modelem dla zmiennych zależnych charakteryzujących się losowością opisywaną rozkładem normalnym oraz dopuszcza ujemne wartości zmiennych zależnych.

W ostatnich latach przyjmuje się najczęściej dwa rozkłady jako odpowiednie do opisu losowości występowania liczby wypadków: rozkład Poissona oraz rozkład Pascala (ujemny rozkład dwumianowy), który w zasadzie jest modyfikacją rozkładu Poissona [2, 3, 4]. Przyjmując losowość zjawiska opisywaną rozkładem Poissona funkcja regresji przyjmuje postać:

$$\lambda = e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)} \quad (2)$$

gdzie λ jest wartością oczekiwaną zmiennej zależnej, a reszta parametrów i zmiennych przyjęta jest jak we wzorze (1).

Przyjęta postać modelu regresyjnego jak we wzorze (2) opisuje prawidłowo zależność pomiędzy zmienną zależną λ a grupa zmiennych niezależnych x_i tylko wtedy, gdy spełniony jest warunek rozkładu zmiennej zależnej. W przypadku rozkładu Poissona warunkiem tym jest jednakowa wartość średniej oraz wariancji opisywanych wielkością λ .

Częstym przypadkiem występującym podczas modelowania liczb zdarzeń drogowych jest nadmierne rozproszenie zmiennej zależnej. Warunek równości wartości średniej i wariancji nie jest wtedy spełniony. Odpowiednim modelem, dopuszczającym nadmierne rozproszenie danych jest model oparty na rozkładzie Pascala, nazywanym również rozkładem ujemnym dwumianowym. Rozkład Pascala jest modyfikacją rozkładu Poissona, w którym parametr λ sam w sobie jest losowy, opisany rozkładem Gamma. Postać funkcji regresji jest wtedy następująca:

$$\lambda = e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \xi)} \quad (3)$$

gdzie φ jest parametrem błędu opisanym rozkładem Gamma, a reszta parametrów i zmiennych przyjęta jest jak we wzorze (2). W przyjętym modelu zależność pomiędzy wartością oczekiwaną λ a wariancją σ jest następująca [4]:

$$\sigma = \lambda(1 + \theta \cdot \lambda) \quad (4)$$

gdzie ν jest parametrem rozproszenia. Jeżeli ν nie różni się istotnie od zera (sprawdzając odpowiednim testem statystycznym), wówczas odpowiednia jest regresja oparta na rozkładzie Poissona.

2. Zebrana baza danych o zdarzeniach drogowych oraz przeprowadzona analiza regresji

W celu przeprowadzenia analizy regresji wielokrotnej, określającej zależność pomiędzy średnią roczną liczbą zdarzeń drogowych LZ a grupą zmiennych niezależnych charakteryzujących ruch drogowy, wybrano odcinek autostrady A-4 przebiegający przez teren województwa dolnośląskiego. Bazę danych o wypadkach drogowych na tym odcinku utworzono na podstawie danych o zdarzeniach z Komendy Wojewódzkiej Straży Pożarnej we Wrocławiu. Informacje dotyczące czynników określających ruch drogowy uzyskano z materiałów Generalnej Dyrekcji Dróg Krajowych i Autostrad. Analizowany odcinek podzielono na sekcje między poszczególnymi zjazdami-wjazdami autostradowymi stosownie do pomiarów średniego dobowego natężenia ruchu drogowego (SDR). Otrzymane w ten sposób pododcinki zestawiono w tabeli 1. Dokumentacja zdarzeń dotyczy okresu z lat 1999-2005. W tym czasie doszło do 1015 zdarzeń drogowych.

Tabela 4. Analizowany odcinek autostrady z podziałem na odpowiednie sekcje

Odcinek	Długość [km]	Opis odcinka	SDR [p/24h]	
			2000	2005
I	10,8	Krzywa-Budziwojów	9201	12820
II	11,1	Budziwojów-Czerwony Kościół	9486	13702
III	8,2	Czerwony Kościół-Legnica	10405	14700
IV	12,5	Legnica-Wądroże.	11601	15754
V	7,5	Wądroże-Budziszów	11601	17772
VI	13,3	Budziszów -Kostomłoty	12377	17878
VII	11,9	Kostomłoty-Kąty Wrocławskie	16977	21293
VIII	15,9	Kąty Wrocławskie -Węzeł Bielany	18468	24027
IX	11,2	Węzeł Bielany -Krajków	13456	22716
X	14,2	Krajków-Brzezimierz	12943	22453
XI	14,7	Brzezimierz-Granica województwa	12818	21767

2.1 Przyjęty model regresyjny oraz wyniki analiz

Podczas modelowania liczby zdarzeń drogowych wykorzystano model regresyjny oparty na przyjęciu losowości zjawiska opisaną modelem Poissona. Każda otrzymana zależność regresyjna sprawdzona została pod kątem nadmiernej dyspersji zmiennej zależnej, którą jest oczekiwana liczba

zdarzeń w roku LZ. W tym celu oszacowano parametr ν oraz sprawdzono jego istotność statystyczną. Jeżeli okazywał się nieistotny statystycznie, wówczas przyjmowano model Poissona za adekwatny. Sprawdzono również współczynniki korelacji wielowymiarowej określających udział wariancji wyjaśnionej modelem za pomocą trzech różnych metod: standardowego współczynnika korelacji R^2 , ważonego współczynnika korelacji R_w^2 oraz współczynnika korelacji R_{ft}^2 , opartego na transformacji statystycznej Freeman'a-Tukey'a. Obliczono również współczynniki korygowane, określające udział wyjaśnionej wariancji systematycznej, z uwagi na przyjęcie rozkładu Poissona jako rozkładu zmiennej zależnej. Obliczono także statystykę MD, określającą istotność przyjętego modelu.

W przyjętym modelu uwzględniono zmienne wyjaśniające według wzoru:

$$\lambda_i = Lod_i^{b_1} \cdot SDR_i^{b_2} \cdot e^{\sum b_j \cdot x_{ji}}, j > 2 \quad (5)$$

gdzie poszczególne wielkości oznaczają:

b – estymowane współczynniki poszczególnych zmiennych,

Lod – długość danej sekcji [km],

SDR – średni dobowy ruch na danej sekcji [p/24h],

x_j – dodatkowe zmienne objaśniające takie jak udział pojazdów ciężkich Uc [%], liczba wszystkich zjazdów-wjazdów autostradowych ZW , liczba zjazdów-wjazdów drogowych ZW_D , liczba zjazdów-wjazdów połączonych z drogami krajowymi ZW_k , gminnymi ZW_g , lokalnymi ZW_l , liczba zjazdów-wjazdów parkingowych ZW_p , modernizacja odcinka PoM (wartości 0 do 1).

Dodatkowo zastosowano transformacje logarytmiczne zmiennych Lod i SDR , co pozwalało na osiągnięcie wartości 0 przez zmienną zależną, jeżeli długość odcinka lub natężenie ruchu były równe 0.

Etapy budowania modelu, poprzez włączanie kolejnych zmiennych objaśniających zestawiono w tabelach 2 i 3. Wyróżnione modele odznaczają się poprawnością przyjęcia założenia rozkładu Poissona przy estymacji parametrów b_i . Odznaczają się nieistotnością parametru ν oraz wysoką wartością wariancji wyjaśnionej modelem. Parametry nieistotne statystycznie zaznaczono kursywą.

Otrzymane modele regresyjne wykorzystać można w prognozowaniu liczby zdarzeń drogowych znając przyszłe wartości zmiennych niezależnych (czynników określających ruch drogowy), następnie otrzymane wielkości uwzględnić przy obliczaniu wskaźników wypadkowych. Należy jednak uwzględnić następujące uwarunkowania podczas działań prognostycznych:

- otrzymane modele regresyjne odnoszą się tylko do badanego odcinka drogowego, ponieważ udział poszczególnych zmiennych niezależnych określony został dla danego miejsca [2],
- trudno przesądzać o stałości oddziaływania na bezpieczeństwo uzyskanych prawidłowości w prognozach wieloletnich,
- należy pamiętać, że zwykle udaje się wyjaśnić część wariancji systematycznej liczby wypadków, poza tym pozostaje jeszcze człon losowej wariancji (w modelu Poissona równy wartości oczekiwanej λ), na który nie można mieć wpływu [1]. Szacowana jest oczekiwana liczba zdarzeń drogowych, najbardziej prawdopodobna ze wszystkich wielkości jakie mogą się pojawić, ale nie jest to wielkość pewna.

Tabela 5. Otrzymane modele regresyjne

Model	Zmienna										
	Stala	Lod	SDR	Uc	ZW	ZWD	ZWk	ZWg	ZWl	ZWp	PoM
1	-7,06	1,65	0,56	-	-	-	-	-	-	-	-
2	-7,23	1,55	0,89	-0,09	-	-	-	-	-	-	-
3	-9,22	1,49	0,82	-	0,01	-	-	-	-	-	-
4	-5,19	2,25	0,60	-0,11	-0,05	-	-	-	-	-	-
5	-6,75	1,68	0,83	-0,09	-	-0,02	-	-	-	-	-
6	-2,33	2,45	0,22	-0,09	-	-	0,05	-0,16	-0,13	-	-
7	-4,56	2,60	0,43	-0,10	-	-	-0,02	-0,08	-0,07	-0,09	-
8	-9,35	2,16	0,90	-0,05	-0,05	-	-	-	-	-	-0,77
9	-9,94	2,15	0,99	-0,06	-	-0,09	-	-	-	-	-0,86
10	-9,83	2,26	0,95	-0,06	-	-	-0,07	-0,10	-0,10	-	-0,85
11	-11,49	2,36	1,12	-0,06	-	-	-0,12	-0,04	-0,06	-0,06	-0,80
12	-9,91	1,59	1,06	-0,05	-	-	-	-	-	-	-0,65

Tabela 6. Istotność statystyczna wybranych modeli

Model	Współczynnik korelacji wielowymiarowej						Dewiacja modelu		Parametr Y (NBM)	
	R ²	Rp ²	Rw ²	Rwp ²	Rft ²	Rftp ²	MD	p	Y	p
8	0,72	0,88	0,69	0,87	0,69	0,87	261,26	10 ⁻⁷	0,034	0,1936
9	0,74	0,90	0,70	0,88	0,70	0,90	267,63	10 ⁻⁷	0,027	0,2617
10	0,76	0,92	0,72	0,91	0,72	0,92	269,76	10 ⁻⁷	0,020	0,3763
11	0,78	0,95	0,75	0,94	0,75	0,96	269,83	10 ⁻⁷	0,011	0,6067
12	0,70	0,85	0,67	0,84	0,67	0,85	248,61	10 ⁻⁷	0,053	0,1054

3. Podsumowanie

Określenie wpływu poszczególnych czynników charakteryzujących ruch drogowy na bezpieczeństwo obiektu jest zadaniem trudnym, wymagającym zebrania odpowiednich danych w dostatecznie długim okresie. Wybór odpowiednich zmiennych niezależnych powinien być poparty teorią i doświadczeniem, a nie tylko koniecznością najlepszego dopasowania modelu regresyjnego do wyników obserwacji zmiennej zależnej. Przyjęta postać funkcji regresji oraz otrzymane zależności dowodzą odpowiedniego skompilowania zmiennych objaśniających. Zachowując zalecane uwarunkowania, otrzymane modele można wykorzystać do oszacowania przyszłych oczekiwanych liczb zdarzeń drogowych oraz ich wskaźników.

Praca naukowa finansowana ze środków na naukę w latach 2007-2008 jako projekt badawczy nr 4564/B/T02/2007/33.

Spis literatury

1. Baruya A.: *Speed-accident relationships on European roads*, Transport Research Laboratory, U.K., 1998.
2. Evans A. W.: *Estimating transport fatality risk from past accident data*, Accident Analysis & Prevention, 35, 2003.
3. Miaou S.P., Lum H.: *Modeling vehicle accidents and highway geometric design relationships*, Accident Analysis & Prevention, 25, 1993.
4. Otto W.: *Ubezpieczenia majątkowe*, WNT, Warszawa, 2004.
5. Washington S.P., Karlaftis M., Mannering F.: *Statistical and econometric methods for transportation data analysis*, CRC, Boca Raton, FL, 2003.



M.Sc. **SIPA Jacek**, PhD student, Wrocław University of Technology, Faculty of Mechanical Engineering, Institute of Machine Design and Operation, specialization: engineering, risk and safety analysis, technical diagnostic.



Dr. Eng. **MLYNCZAK Marek**, Wrocław University of Technology, Faculty of Mechanical Engineering, Institute of Machine Design and Operation, specialization: technical systems reliability, risk and safety analysis, discrete simulation.

