Journal of KONBiN 4(24) 2012
ISSN 1895-8281
DOI 10.2478/jok-2013-0060

THE MULTIVARIATE MULTILEVEL ANALYSIS OF DIFFERENT REGIONAL FACTORS IMPACT ON ROAD SAFETY IN EUROPEAN COUNTRY REGIONS

WIELOPOZIOMOWA ANALIZA WIELOCZYNNIKOWA WPŁYWU RÓŻNYCH CHARAKTERYSTYK REGIONALNYCH NA POZIOM BEZPIECZEŃSTWO RUCHU W REGIONACH KRAJÓW EUROPEJSKICH

Joanna Wachnicka, Leszek Smolarek

Politechnika Gdańska, Akademia Morska
joanna.wachnicka@wp.pl, leszsmol@am.gdynia.pl

***Abstract**: In this research, the effect of the European regional differences on the risk of been killed in road accident in these regions is investigated. Factors which differentiate regions can be described as automotive, economic, infrastructural, demographic, geographic. Analyzed risk, counted as regional fatality rate was modeled by the Poisson model. Because of regional diversity of Poisson distribution parameters, the Poisson distribution was adopted. The multivariate multilevel model was developed, in order to asses combined effects of different factors on road safety level in regions. Such models could be helpful to forecast such risk on the regional level and focus safety actions precisely for region, because the national actions and their results differentiate in regions so far.*

***Keywords:** road safety, hierarchical modeling, fatality rate, regions,*

***Streszczenie:** W artykule przedstawiono wpływ cech różnicujących regiony krajów europejskich na ryzyko bycia ofiarą śmiertelną wypadku drogowego. Zmienne różnicujące regiony to m.in. charakterystyki motoryzacyjne, ekonomiczne, infrastrukturalne, demograficzne i geograficzne. Założono, że ryzyko mierzone jako wskaźnik zabitych w regionach ma rozkład Poissona, a w procesie modelowania przyjęto uogólniony model Poissona. Opracowano wielopoziomowy model wieloczynnikowy w celu oceny złożonego, wzajemnego wpływu różnych czynników na poziom bezpieczeństwa w regionach, który może pomóc w prognozowaniu analizowanego wskaźnika na poziomie regionalnym oraz ukierunkowaniu działań służących poprawie bezpieczeństwa w danym regionie.*

***Słowa kluczowe:** bezpieczeństwo ruchu drogowego, modelowanie hierarchiczne, wskaźnik zabitych, regiony,*

1. Analyses of regional data - hitherto approach

The issue of road traffic safety has been studied on international level for many years now and many publications have appeared on the issue. It is a problem, which embraces many scientific fields seemingly not related to each other, such as road engineering, economy, mathematics, transport or medicine. Studies of changes occurring in traffic safety in respective countries were often based on analyses of changes in road fatality rate in a given time horizon [1],[2]. Sometimes researchers analysed, which factors may influence observed changes [3]. Such kind of approach could be a source of information for legislative authorities or road administrators to be used for effective road traffic safety management.
However, as experience shows, the actions run at the national level translate into effects in lower administration units to variable extents. There might be an improvement of traffic safety observed in one region and none in the other. Therefore it is justified to analyse national as well as regional characteristics of a given area and their potential impact on the road safety level.
In the researches made at the national level gross national product per capita [4] and transport activity [5] are often listed as significant factors influencing traffic safety. Since transport activity figure is frequently unavailable in regional databases, in their papers concerning regions the researchers pointed to population density as independent variable significantly influencing traffic safety level in regions, replacing other characteristics unavailable at this aggregation level [6]. Multivariate studies showed the impact of such factors as road density [7]. As the literature study indicates the researchers focusing on the analyses of road safety level in the regions have not taken into account national characteristics in regional analyses, and such approach has been applied for this paper.

2. Multilevel model – general principles

In the article an attempt to combine national and regional characteristic in a single model has been presented. For this purpose the data on fatality rate in a given region have been collected as well as additional characteristics describing regions and countries. Such approach was inspired by the fact that there are good characteristics available on national level that can efficiently differentiate safety level in the regions of respective countries, but unfortunately unavailable on regional aggregation level. On the other hand respective regions of a given country differ one from another for instance in respect of population density or road network density and these elements are worth considering in the model. Since a significant dispersion of a number of fatalities was observed, the decision was taken to model road fatality rate in relation do demography calculated as a number of fatalities per one hundred thousand inhabitants. The model is meant to take the following formula:

$$FATALR = \alpha \cdot MODEL_{NAT.}^{\beta_1} \cdot MODEL_{REG.}^{\beta_2} \cdot NPPC_{NAT}^{\beta_3} \quad (1)$$

where:
FATALR - *the road fatality rate in relation to demography in a given region [fatalities/100 thou. inhabitants]*
$MODEL_{NAT.}$ - *model for national data*
$MODEL_{REG}$ - *model for regional data*
$NPPC_{NAT.}$ - *model describing changes in average national product per capita*
$\alpha, \beta_1, \beta_2, \beta_3$ - *estimated parameters*

FATALR estimation is based on the assumption that the parameter has Poisson distribution. National models were created on the basis of data coming from 11 European countries, whereas in case of regional models the focus in this paper has been on two European countries differing substantially in terms of road safety: Sweden being a role model country for European states in terms of actions for the improvement of road safety and Czech Republic, where the average fatality rate is double the Swedish figure, probably due to cultural, political and economic differences. Histograms of FATALR values in the regions of the compared countries within analysed period of 1999-2008 presented on Figure 1 indicate that there is no ground to dismiss the hypothesis of Poisson distribution, which is frequently adopted in the analyses of safety level [8]. In further analyses, according to this assumption, the FATALR index will be alternatively referred to as λ.

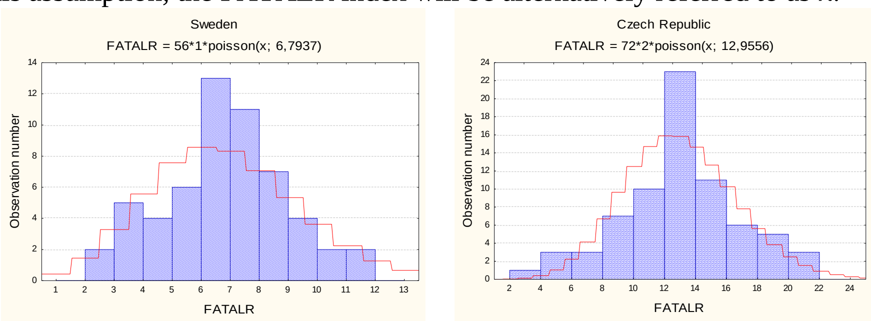

Fig. 1. Histograms of analysed FATALR rates in regions of Sweden and Czech Republic.

3. Creation of the model

Taking into account the adopted assumptions, parameter λ was evaluated with 95% probability for the data set from the compared countries. Then the set of independent variables was created, available only at the national level of aggregation, characterizing the country, used for the purpose of creating descriptive models for estimated λ parameters ($MODEL_{NAT.}$). In this paper the influence of the following factors was analysed: human development index - HDI, percentage of passenger cars older than 10 years in total fleet of the country -OLD as well as inflation level -INFL, calculated as averages from 10 years.

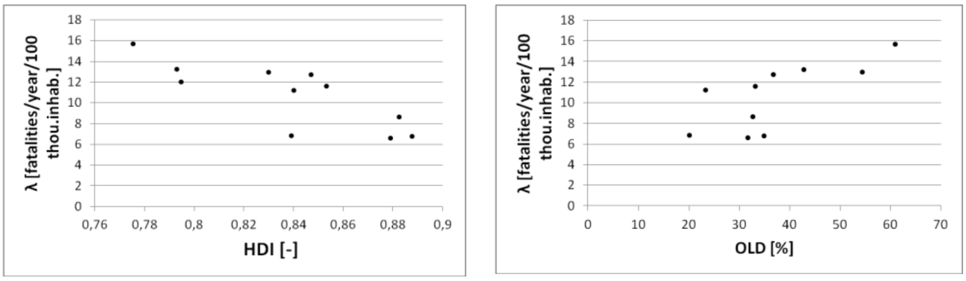

Fig. 2. Chart of dependence of λ on: HDI value in a given country, and the percentage of passenger cars older than 10 years old.

As shown by Figure 2 these rates may influence the studied parameter. A trend has been observed of λ parameter reduction with the increase of human development index as well as the inverse relationship in case of percentage of old passenger cars. Power-exponential model has been prepared:

$$\lambda = \alpha \cdot x_1^{\gamma} \cdot e^{(\beta_1 \cdot x_2 + \beta_2 \cdot x_3)} \tag{2}$$

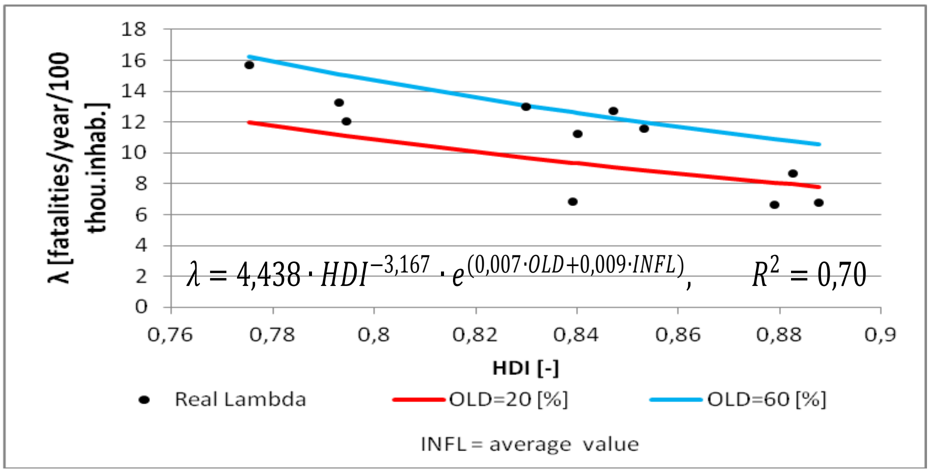

Fig. 3. Illustration of power-exponential model taking into account human development index, percentage of passenger cars older than 10 years in total vehicle fleet and the inflation level.

The model's Q factor is 0.7. λ parameter decreases with the growth of human development index HDI, though the λ parameter will be higher if the rate of old cars in the total fleet of passenger cars is high.

The next step was to create models describing impact of regional characteristics in respective countries on FATALR values in these regions. For this purpose the separated regional database for each country was created and efforts were made to develop the model of impact of individual variables on modelled dependent variable. For all countries one type of model was checked, initially developed on the basis of total database of all European regions. General shape of this model has been presented below (3), whereas in Table 1 calculated coefficients in the model have been presented country-wise. Cluster analysis allowed determining classes of interdependent variables. In respective models the impact of individual classes was considered by choosing their representatives.

$$MODEL_{REG} = \alpha \cdot (lnNPCC)^{\beta} \cdot e^{(\gamma_1 \cdot lnDPOP + \gamma_2 \cdot lnVEHD + \gamma_3 \cdot ROADC + \gamma_4 \cdot STUD + \gamma_5 \cdot MOTORP)} \quad (3)$$

where:

lnNPCC - *logarithm of national product per capita [-]*
lnDPOP - *logarithm of density of population [-]*
lnVEHD - *logarithm of total vehicle density [-]*
lnROADC - *model describing changes in average national product per capita*
STUD - *percentage of secondary school and university students [%]*
MOTORP - *percentage of motorways [%]*
α, β, γ_i, - *estimated parameters*

Table 1. List of parameters of regional models for selected countries

Model	**R^2**	**α**	**β**	**γ_1**	**γ_2**	**γ_3**	**γ_4**	**γ_5**
Czech Rep. 1	0,81	35,529	-2,259		0,174	1,056		
Czech Rep. 2	0,83	59,816	-2,048		0,166	0,878	-0,030	
Czech Rep. 3	0,85	307,415	-2,751			0,688	-0,038	0,051
Sweden 1	0,65	482,858	-3,229	-0,116				
Sweden 2	0,71	937,868	-4,291			0,142		

Global national and local models and local regional models are models, in which the majority of independent variables are characterised by slight variation in time. The only variable that changes in time dynamically and present in almost all models is national gross product per capita. Analysis of variation of average NPPC in time showed that the analysed countries are characterised by two types of trends of NPPC variation in time. First type of trend is linear, typical of the “Old Europe” countries, where the economic situation is stable and living standard is rising. The second one is nonlinear trend, probably stemming from the dynamic changes occurring in these countries after their access to the European Union, fig. 4.

After development of partial models the model composed of all partial models was created with the final formula as follows:

$$FATALR = 10{,}171 \cdot MODEL_{NAT.}^{0{,}240} \cdot MODEL_{REG.}^{0{,}088} \cdot NPPC_{NAT.}^{-0{,}28} \quad (4)$$

It is a multiplicative model, the components of which have Q factors ranging between 0.99 and 0.7, whereas resultant model has Q factor equal to 0.41. This is a result of using annual average values as input data, in order to eliminate temporary variations that could conceal the character of respective impacts. At the current, initial stage of studies such result might be considered satisfactory.

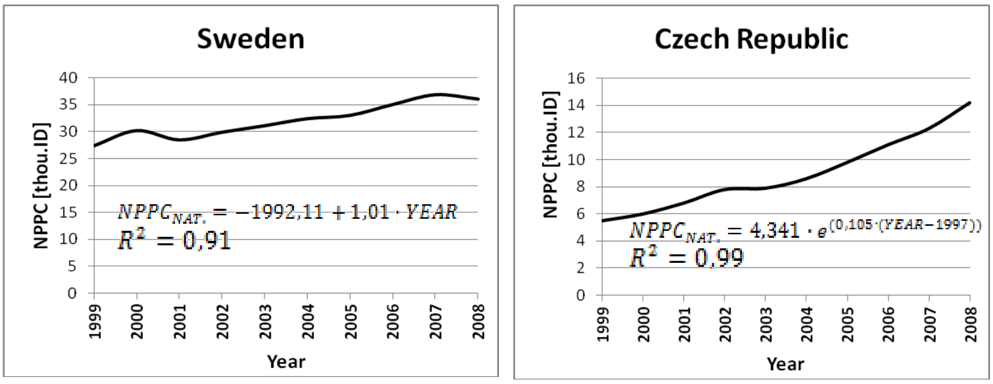

Fig. 4. Chart of changes in average gross national product per capita in the years 1999-2008 in Sweden and Czech Republic.

4. Model sensitivity analysis

Sensitivity analysis is based on prediction of model results, while applying variable factor arrangements having impact on results, i.e., for instance, checking how the input variable changes when we change one of the input variables by 1%. It represents a useful tool for risk reduction. Sensitivity analysis allows calculating border level of values of respective factors, which ensure reaching specified level of safety. It may be used for considering potential impact of different factors, which may be related to different local conditionings.
Sensitivity of the model defined by formula (3), is determined by elasticity of function calculated for variable x_r according to following the formula

$$\varepsilon_r^f = \frac{\partial MODEL_{REG}}{\partial x_r} \cdot \frac{x_r}{MODEL_{REG}}$$

where:

$$x_r \in \{NPPC, DPOP, VEHD, ROADC, STUD, MOTORP\}.$$

When calculating for respective variables we receive

$$\varepsilon_{NPPC}^f = \frac{\beta}{\ln NPPC}\,;\ \varepsilon_{DPOP}^f = \gamma_1 DPOP\,;\ \varepsilon_{VEHD}^f = \gamma_2 VEHD\,;\ \varepsilon_{ROADC}^f = \gamma_3 ROADC\,;\ \varepsilon_{STUD}^f = \gamma_4 STUD\,;\ \varepsilon_{MOTORP}^f = \gamma_5 MOTORP$$

Model allows determining what rate it is dependent on, but not how it is dependent, in particular determining what it is dependent on more strongly and what to a small extent. Results of sensitivity analysis allow hierarchization of factors and variables in respect of the impact “strength”. Thus, each time one of the input variables is changed, and the remaining ones have original, initial values. The conclusions are simple, which is reflected in elasticity presented in Table 2.

Table 2. List of sensitivity parameters for regional models from Table 1

Model	εNPPC	εDPOP	εVEHD	εROADC	εSTUD	εMOTORP
Czech Rep. 1	21,44067		0,174	0,7392		
Czech Rep. 2	19,43802		0,166	0,6146	-0,768	
Czech Rep. 3	26,11035			0,4816	-0,9728	0,0408
Sweden 1	-4,65846	-0,116				
Sweden 2	-6,1906			0,284		

In the case of model given by the formula (4) we receive $\varepsilon MODEL_{NAT.} = 0.24$; $\varepsilon MODEL_{REG.} = 0.088$, $\varepsilon NPPC_{NAT.}$ -0.28. It stems from the fact that the partial elasticity of power function is constant and is equal to the value of power of each of the variables.

5. Conclusion

The conducted heuristic statistical analysis, the results of which were presented in previous sections, indicates usefulness of employing local characteristics for evaluation of road transport safety. Due to varying availability of registered data, both at national and regional level, the cluster analysis was necessary in order to single out classes of “correlated” variables of which each one could be used in the model as representative of specified group of effects (impacts). Division into impact groups was conducted both at the level of countries and regions. All variables present in the models are significant. NPPC variable has the biggest impact.

Acknowledgement

The paper has been funded by grant-in-aid for carrying out scientific research or development work and related tasks, aimed at the progress of young scientists and doctoral students.

6. References,

[1] Commandeur J. J. F., Bijleveld F. D., Bergel-Hayat R., Antoniou C., Yannis G., and Papadimitriou E., “On statistical inference in time series analysis of the evolution of road safety.,” *Accident; Anal. Prev.*, pp. 1–11, Dec. 2012.

[2] Holló P., Eksler V., and Zukowska J., “Road safety performance indicators and their explanatory value: A critical view based on the experience of Central European countries,” *Safety Science*, vol. 48, no. 9, pp. 1142–1150, Nov. 2010.

[3] K. Jamroz, “Metoda zarządzania ryzykiem w inżynierii drogowej,” 2011.

[4] Van Beeck A. E., Mackenbach E., , Looman J. P., , & Kunst C. W., , “Determinants of Traffic Accident Mortality in the Netherlands: A

Geographical Analysis.," *International Journal of Epidemiology,* vol. 20, no. 3, pp. 698–706, 1991.
[5] K. Jamroz, "The impact of road network structure and mobility on the national traffic fatality rate," *Energy Efficient Transportation Networks, International Scientific Conference,* Paris 2012.
[6] Lassarre S. and Thomas I., "Exploring road mortality ratios in Europe : national," *J.R. Statist.Soc.A* ,pp. 127–144, 2005.
[7] Erdogan S., "Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey.," *Journal of safety research,* vol. 40, no. 5, pp. 341–51, Oct. 2009.
[8] Hinkley, D.V.;Reid, N." Statistical Theory and Modelling: In Honour of Sir David Cox Frs", Chapman & Hall, 1991

M.Sc. Eng. Joanna Wachnicka*, Gdańsk University of Technology, assistant in the Department of Highway Engineering. She specializes in road and transport engineering as well as road traffic safety. She has published 12 articles in journals and conference proceedings.*

dr hab. Leszek Smolarek*, Gdynia Maritime University, Gdynia, Professor at the Mathematics Department. He has 25 years experience in teaching and research work. He has published more than 30 books, reports and papers in journals and conference proceedings.*