

# Modelling fatalities on regional road networks

Joanna Wachnicka<sup>1,\*</sup>

<sup>1</sup>Gdansk University of Technology, Faculty of Civil and Environmental Engineering, Gdansk, Poland

**Abstract.** During the last decade Poland's road fatalities went down by 44%. The trend differs from region to region. Effective road safety management in regions requires tools for forecasting road safety measures and identifying factors influencing road fatality numbers. Mathematical models can provide such tools. They take into account local characteristics such as: demography, economy, infrastructure and motorization. Such models could be used for better regional road safety management. This paper presents an attempt to build such models used for forecasting road fatalities in EU regions and Polish voivodships.

## 1 Introduction

Road accidents worldwide claim nearly 1.3 m lives annually causing an economic loss that represents 2 - 3% of the world's GDP. Sadly Poland ranks among the top thirty countries in the world and the European Union's top five for its road traffic risk. Despite the EU being the world leader on road safety, four countries: Germany, France, Italy and Poland contribute to more than 50% of the EU's road fatalities. The risk of becoming a fatality in relation to Poland's population is double the European Union's average and as much as three times more than in the United Kingdom, the Netherlands and Sweden.

It is estimated that until 2030 road fatalities may double if the worldwide community fails to take effective and comprehensive steps to improve road safety. Being involved in a road accident may become one of the five most frequent causes of death of the globe's population. This is what spurs a variety of government organisations and NGOs into action geared towards reducing road accidents. The United Nations with its World Health Organisation (WHO) have prepared recommendations for the countries, designed to improve road safety. They are part of the Decade of Action for Road Safety adopted in 2011 (for the years 2011 – 2020). The fundamental intention behind this idea is to save 5 million people from death in a road accident in the years 2011 – 2020. This is to be achieved by efforts delivered by the particular countries, including the development and implementation of road safety programmes, appointing road safety management bodies, effective accident data collection and increasing spending on road infrastructure to ensure that all road projects comply with safety requirements when they are designed, built and maintained [1].

Poland has its own systemic solutions designed to improve road safety. The country's first road safety programme (adopted by the government) was developed in 1972 [2]. In 1993 the National Road Safety Council was established at the central level with Regional Road Safety Councils appointed at the regional level. A year later the minister of transport asked the Committee of Scientific Research to conduct a research project called "Integrated Road Safety Programme" under the acronym of GAMBIT. Developed by an interdisciplinary team

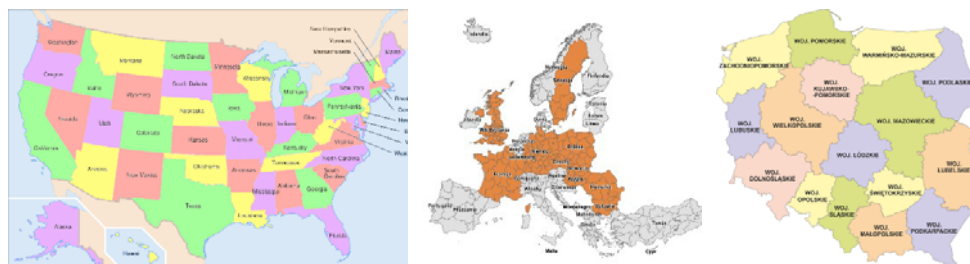
---

\* Corresponding author: [joanna.wachnicka@wilis.pg.gda.pl](mailto:joanna.wachnicka@wilis.pg.gda.pl)

of experts headed by professor R. Krystek, the programme set quantitative targets to be attained within the next 5 years and the projects to be implemented within this time. Two years into the programme it was established that the targets cannot be reached. More GAMBIT road safety programmes followed in 2000, 2005 and in 2012. While they were nation-wide programmes, new programmes were designed specifically for the regions and cities to ensure that action at the lower levels can have a stronger effect on the regional problems. The regional or city road safety forecasts which they used were largely based on police statistics and earlier domestic and international experience. What they lacked, however, were mathematical tools to support road safety forecasts such as mathematical models that take account of regional characteristics and their effect on the regional situation. Because no country has solved the problem of modelling regional road safety, the decisions were made to build such road safety models for regions in the United States, Europe and Poland. Such models, due to the inclusion of multiple factors, may be helpful in analysing the impact of the individual on the modelled variable, thus helping to better manage road safety in the regions.

## 2 Characteristics of regional road safety

Road safety on the US road network was analysed using data from 51 states. In the case of EU regions, data were collected from 188 NUTS 2 regions of the following countries: the United Kingdom, Czech Republic, Germany, Poland, France, Spain, Slovakia, Hungary, Sweden, Bulgaria and Romania. In Poland, the analysis looked at 16 regions (voivodeships) (Fig. 1). The time period under analysis in all the cases spanned from 1999 to 2008 with Polish regions as the exception because the data covered the years 1999-2014. Detailed analyses of the levels of road safety in groups of regions showed significant differences between them.

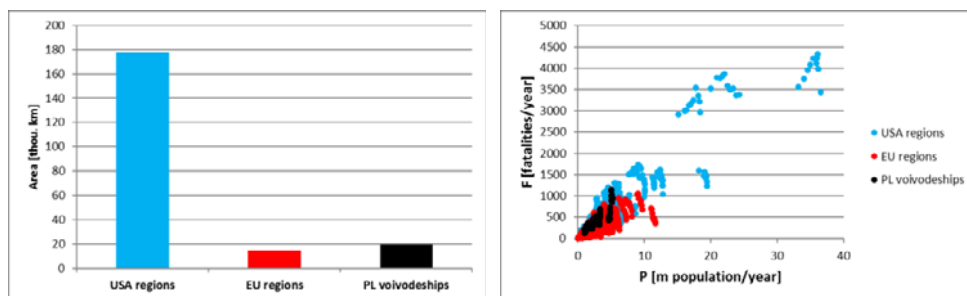


**Fig. 1.** Analysed regions in particular data groups.

US regions showed strong regional differentiations. As an example, the population in the states ranges from 480,000 to more than 36 m people and the area is between 157,000 km<sup>2</sup> to app. 1.5 m km<sup>2</sup>. Population density shows strong variations as well with values ranging from 420 people/km<sup>2</sup> to more than 5,500 people/km<sup>2</sup>. Road accident fatalities show a range between 34 to app. 4,300 people annually and about 800 people on average annually. The population of the EU regions studied is between 70,000 to 11.7 m and the area is from 100 km<sup>2</sup> to 94,000 km<sup>2</sup>. Population density varies from 5 people/km<sup>2</sup> to more than 5,500 people/km<sup>2</sup>. Road accident fatalities are within a range of 1 to almost 1,150 people, with the average of 213 people. The area of Polish voivodeships ranges from app. 9.4 to 35,600 km<sup>2</sup>, population from app. 1 do 5.3 m, and the number of road fatalities regionally varies from 81 to more than 1,100 people annually.

The difference in the average area of regions in the United States, Europe and Poland is shown in Fig. 2. The figure also presents the relation between fatalities and the region's

population. As we can see, European and Polish regions are similar while the American states differ significantly from the two groups. The average gross domestic product GDPPC in the US is higher than Europe's by 25% and almost three times higher than a Polish voivodeship's GDPPC. At the same time the average fatalities on the roads of the particular US states are almost four times higher than in Europe and twice as high as that in Poland.



**Fig. 2.** Analysed regions in particular data groups.

**Table 1.** Average, minimal and maximal values for the basic characteristics in the three analysed databases.

Data base	Characteristics	average	minimum	maximum
States AP	Area [thou.km <sup>2</sup> ]	177.572	0.157	1464.195
	Population [m]	5.712	0.480	36.580
	Motorization rate [veh./person]	0.862	0.253	1.512
	Gross Domestic Product [thou. ID/person]	41.200	25.039	151.392
	Fatalities [people]	812	34	4329
	Fatality rate [people/100 thou.mk.]	16.134	1.032	39.408
Regions EU	Area [thou.km <sup>2</sup> ]	14.486	0.32	154.312
	Population [m]	2.081	0.119	11.694
	Motorization rate [veh./person]	0.535	0.098	1.603
	Gross Domestic Product [thou. ID/person]	30.183	5.686	116.83
	Fatalities [people]	162	6	1057
	Fatality rate [people/100 thou.mk.]	9.861	1.055	25.839
Voivodeships PL	Area [thou.km <sup>2</sup> ]	19.524	9.412	6.645
	Population [m]	2.396	1.008	5.204
	Motorization rate [veh./person]	0.357	0.166	0.588
	Gross Domestic Product [thou. ID/person]	12.069	7.932	25.951
	Fatalities [people]	367	128	1146
	Fatality rate [people/100 thou.mk.]	15.674	9.145	24.554

Preliminary analyses have shown that regional data cannot be analysed in one set. This is because the scale of the regions differs significantly, as does variable availability and the scope of change in the analysed risk measures. As a consequence, the analyses that followed were conducted separately for each group of regions.

A number of independent variables were collected to build the mathematical models. Divided into groups of factors, the variables are region-specific. The following are the groups of factors:

- demographic (characteristic of the region's population),
- geographic (characteristic of the area and land use),



- motorization (characteristic of the fleet of vehicles using the region's road network),
- mobility (characteristic of the mobility of the population and their transport behaviour),
- infrastructure (characteristic of the structure of the region's road network),
- economic (characteristic of the region's economic development and its investment spending),
- social (characteristic of the region's population in terms of its development, health, behaviour).

Data availability differed from group to group and amounted to 49 independent variables in US regions, 27 in European regions and 71 in Polish regions. While preliminary analyses of the factors are presented in the reviewed articles [3,4], the focus of this article is to describe the models developed for each of the three groups of regions.

### 3 Modelling methods

Preliminary data analyses [3,5] made it clear that modelling fatality numbers was going to be difficult. As a result, the decision was made to create models of the relative fatality rate, the *RFR*. Where:

$$RFR_i = \frac{F_i}{P_i} \quad (1)$$

where:

- $RFR_i$  – fatality rate in the region  $i$  [fatalities /100 thou. people/year]
- $F_i$  – fatalities in the region  $i$  [fatalities /year]
- $P_i$  – population of the region  $i$  [fatalities /100 thou. people/year].

This approach helps to build reliable models and, using the region's forecasted population, calculate the forecasted number of fatalities in relation to specific factors using this formula:

$$F_i = RFR_i \cdot P_i \quad (2)$$

To understand which independent variables may have an effect on the RFR, for each database an analysis was made into the most probable functions which could describe the relations under study.

The following functions were tested: linear, logarithmic, Kuznets, exponential, power and power-exponential. As a result, it was possible to identify the link functions in the new models. The exponential and power-exponential functions were found to work best in the case.

### 4 Results of modelling

In the case of states in the US, the relative fatality rate was strongly influenced by: length of vehicle kilometres travelled per one inhabitant *VKTPC* and income per capita in that area *INPC*. Thanks to variable availability, the models reached a quality factor of app. 75%. Formula 3 shows an example of one of the models.

$$RFR = 14.79 \cdot \exp(-0.03 \cdot INPC + 0.04 \cdot DPR_{(D)} + 0.09 \cdot VKTPC + 0.11 \cdot DME_{(D)}) \quad (3)$$

where:

- *INPC* – national income per capita [thou.ID/person],
- $DPR_{(D)}$  – demographic density of paved roads [thou.km/100 thou. person],
- *VKTPC* – kilometres travelled by vehicles per capita [thou.km/person],

-  $DME_{(D)}$  – demographic density of motorways and express roads [thou.km/100 thou. person].

Analysis of coefficients for the particular variables shows that as national income per capita  $INPC$  grows, the relative fatality rate on the state's roads goes down. At the same time, an increase in vehicle kilometres travelled per capita, demographic density of paved roads and demographic density of motorways and express roads increases the risk of becoming a road accident fatality.

When the  $RFR$  was modelled for European regions, power and exponential functions were used. The level of socio-economic development expressed as gross domestic product per capita  $GDPPC$  turned out to be positively correlated. In the case of vehicle density  $DV$ , it was observed that as the rate begins to grow, so does the relative fatality rate and once it reaches maximum it begins to decrease.

$$RFR = 13.44 \cdot DV^{-0.84} \cdot \exp(0.01 \cdot GDPPC - 0.03 \cdot DR_{(A)} + 2.28 \cdot DV) \quad (4)$$

where:

- $DV$  – total vehicle density [veh./km<sup>2</sup>/year],
- $GDPPC$  – gross domestic product per capita [thou.ID/person/year],
- $DR_{(A)}$  – total road density [km/km<sup>2</sup>/year].

Unfortunately, none of the European regional models exceeded a quality factor of 45%, which is probably due to how strongly differentiated the regions are. Another possible reason may be the short time interval of the analysis. The problem requires further analyses as the work continues.

Because Poland is struggling with its significant fatality rates, the next stage of the work focussed on models of the relative fatality rate in Polish regions. Thanks to detailed data availability, the models use a bigger number of explanatory variables. The resulting model achieved a quality factor of  $R^2=60\%$ . In the case of Polish data, detailed preliminary analyses showed that being a capital region matters. In addition, the relative fatality rate in Polish voivodeships decreases as gross domestic product per capita and total vehicle density increase. Regional road safety improves when regional roads are modernised and more people live in cities.

$$RFR = 24,07 \cdot S^{0,23} \cdot GDPPC^{-0,42} \cdot \exp(-1,02 \cdot DV + 0,33 \cdot DR_{(D)} + 0,51 \cdot PUP - 0,75 \cdot REXV + 0,04 \cdot PUA - 0,07 \cdot PUP + 29,27 \cdot DME_{(D)}) \quad (5)$$

where:

- $GDPPC$  – gross domestic product per capita [thou.ID/person/year],
- $S$  – variable denoting the capital voivodeship [capital=10, others=1],
- $DV$  – total vehicle density [veh./km<sup>2</sup>/year],
- $DR_{(D)}$  – total demographic road density [thou.km/100 thou. person/year],
- $PUP$  – share of urban population [%],
- $REXV$  – spending on regional roads [m PLN/km/year],
- $PUA$  – share of built-up and urbanised land in overall area [%],
- $DME_{(D)}$  – demographic density of motorways and express roads [thou. km/100 thou. person/year].

## 5 Discussion of the results

The models' quality factors varied, depending on data availability and the scope of the regions being modelled. A comparison of the results shows that it is important to collect and share data about the number of kilometres [6,7] travelled on the region's road network. This factor is critical for understanding risk exposure. Collecting data about road expenditure also

turned out to be important [8,9]. This may show indirectly what changes are made to the road network to improve road safety. The results of the modelling confirmed the conclusions from previous research which looked at the effect of regional economic growth on road safety [10,11]. The results of the modelling using three groups of regions suggest that it is difficult to build models to give sufficient explanation of the relations, if the regions are highly differentiated and detailed variables are not available.

## 6 Conclusions

There are many reasons why modelling fatalities is a difficult task. First, if the aim is to build mathematical models that can sufficiently explain the relations, the dependent variable must be as compact as possible. When groups of regions vary significantly in terms of their size, population and, often as a result, fatalities, it is advisable to normalise the data by introducing a relative variable just as was the case in this article. Second, it is important to collect independent variables carefully and with some differentiation to obtain a description of the region's mobility, infrastructure, economy and social aspects. It is also advisable to have expert knowledge of the region to ensure that qualitative variables can be added during modelling to differentiate it from other regions.

## References

1. L. Michalski, K. Jamroz, D. Gajda, Wstępna ewaluacja interim Krajowego Programu Bezpieczeństwa Ruchu Drogowego GAMBIT 2005, in: IX Int. Road Saf. Semin. GAMBIT 2012, pp. 20–21 (Gdansk, 2012)
2. Resolution of the Council of Ministers No. 155/72 of 9 June 1972 on directions and measures intended to improve road safety (Poland, 1972)
3. J. Wachnicka, Modelling selected road safety measures at the regional level in Europe, J. Polish Saf. Reliab. Assoc. **3** pp. 251–258 (2012)
4. J. Wachnicka, K. Jamroz, Identification and analysis of factors influencing road safety in regions, in: Road Saf. Simul. Int. Conf. RSS2013, (Rome, 2013)
5. J. Wachnicka, Identification and comparative analysis of factors influencing road safety in US regions and in Polish voivodships, Transp. Probl. **8** pp. 53–66 (2013)
6. L.S. Robertson, Reducing death on the road: the effects of minimum safety standards, publicized crash tests, seat belts, and alcohol., Am. J. Public Health. **86** pp. 31–34 (1996)
7. G. Scott, G. Pittams, N. Derby, Regression models of New Zealand road casualty data: results of a preliminary investigation, in: Australas. Transp. Res. Forum, pp. 829–854 (Wellington, 1987)
8. S. Gaca, W. Suchorzewski, M. Tracz, *Inżynieria ruchu drogowego. Teoria i praktyka (Road traffic engineering. Theory and practice)* (WKŁ, Warsaw, 2008)
9. J. Kempa, Najczęstsze wady infrastruktury drogowej, in: Bezpieczeństwo Ruchu Miej., pp. 239–287 (WKŁ, Warszawa, 2005)
10. J.Y. Douglas, T.W. Likens, Alcohol Regulation and Auto Fatalities, Int. Rev. Law Econ. **20** pp. 107–126 (2000). doi:10.1016/S0144-8188(00)00023-5
11. S. Hakim, et al., A Critical Review of Macro Models for Road Accidents, Accid. Anal. Prev. **23** pp. 379–400 (1991)