Veronika Valentová
Centrum dopravního výzkumu, v.v.i.
veronika.valentova@cdv.cz

Differences between the expected numbers of accidents for various types of intersection arrangements in Prague

1) INTRODUCTION

Road accidents in the Czech Republic have been a social problem for many years as in other European countries. According to White paper published in 2002 (European Commission, 2001), much has been done in road safety. European Directive 2008/96/ES on Road Infrastructure Safety Management (European Parliament, 2008) was adopted and implemented in national laws and strategies. Much has been done for road safety improvement; however specified goals were not achieved. Identification of hazardous road locations and the influence of different types of arrangements on safety is way how to make the worst locations safer.

Traditional approaches, based on the numeric definition for identification of hazardous road locations and hazardous intersections, are still used. However, these approaches have not been successful enough. Identified locations in different years or periods change a lot. Therefore, new approaches based on the safety performance function are dealt with.
Despite this, road safety modeling is still in the beginning. In the last years in CDV, the first statistical model was made for roundabouts. The model was simple and the empirical Bayes approaches described by Hauer et al. (2002) were not used. The state of the art methods have been used in project IDEKO, which deals with road segments for rural roads of regional importance (Striegler et al., 2012).

In the year 2012, almost 700 people were killed in the Czech Republic in road traffic accidents, approximately one third of them in urban areas. In addition in urban areas 15 000 people were injured. The majority of accidents in towns and cities occur at intersections. For that reason, countermeasures in intersection area are used. The influence of these countermeasures is generally known, although much scope for improvement is in objective evaluation. This study attempts to find some relationship between the different types of intersections and accident frequency.

2) VARIABLES

The model described in this study is for intersections in Prague, the capital city of the Czech Republic. The data used in the model were collected by the Technical Administration of Roads of the City of Prague (TSK Prague). These data were supplemented by police accident records and by own data collection.

The dependent variable was accident frequency. The independent variables, included at first, were the exposure data and the data about intersection characteristics. These data are generally believed to influence road safety. Each variable is described below.
There were more than 400 recorded nodes in Prague. Unfortunately, high number of them had to be excluded for various reasons. At first, a high number of them were not intersections at all. In TSK records there are nodes of roads, not intersections. All of the nodes were checked manually. In many cases, one-way streets were split into higher number of streets. Some of nodes are interchanges and so on. Finally only 194 intersections were left. Descriptive statistics of all available variables are showed in Table 1.

Accident data

Accident data for two years 2009 and 2010 were obtained from the Czech Traffic Police. They include all accidents – injury accidents (accidents with slight, severe and fatal injuries) and property damage only accidents. Underreporting of injury accidents with motor vehicles is minimal; however this is not the case with property damage only accidents. Reporting rate is determined by the reporting threshold. Therefore, only accidents with property damage over 100 000 CZK (4 000 Euro) or accidents in which participants cannot make a deal about who caused the accident are in the police records. For the model, all accidents regardless of their severity were used. Accident data distribution in the year 2009 and 2010 is displayed on graphs in Figure 1. These graphs show the number of intersections with the specific number of accidents.

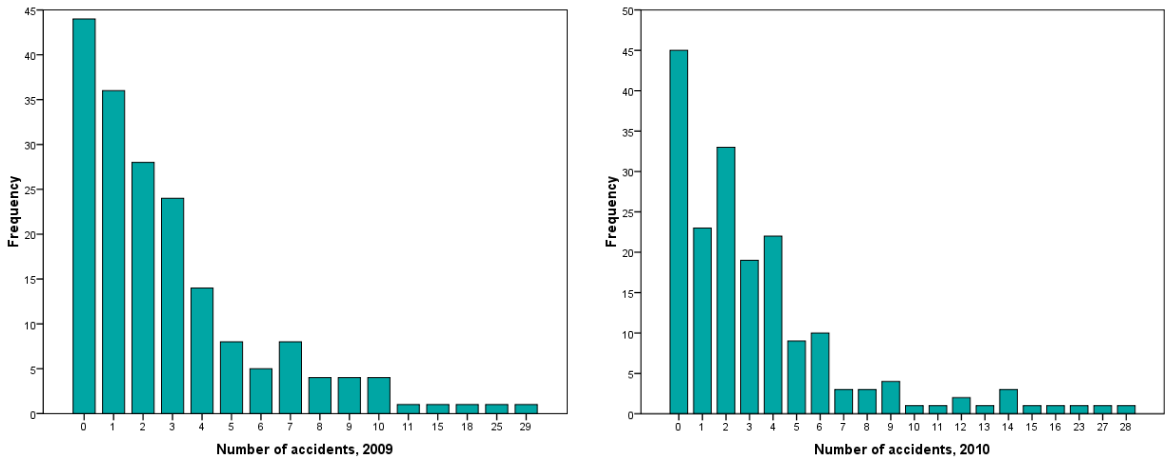


Figure 1 Accident data distribution, years 2009, 2010

Annual average daily traffic

TSK has been making records of AADT on selected roads in Prague. This network covers all roads which are important for traffic in the area. Some of intersections are in the central part of the city, other are in residential, shopping or industrial zones. Traffic volume changes were between approximately 1000 vehicles per day to 85 000 vehicles per day. Histogram of this data was made and intersections with more than 30 000 vehicles per day on the major road were excluded. AADT was divided into intervals and the frequency of the interval occurrence is showed in Figure 2. Value of AADT 30 000 cars a day make up about 95 % of all values. Higher values are rare and are spread in wide range of volume.

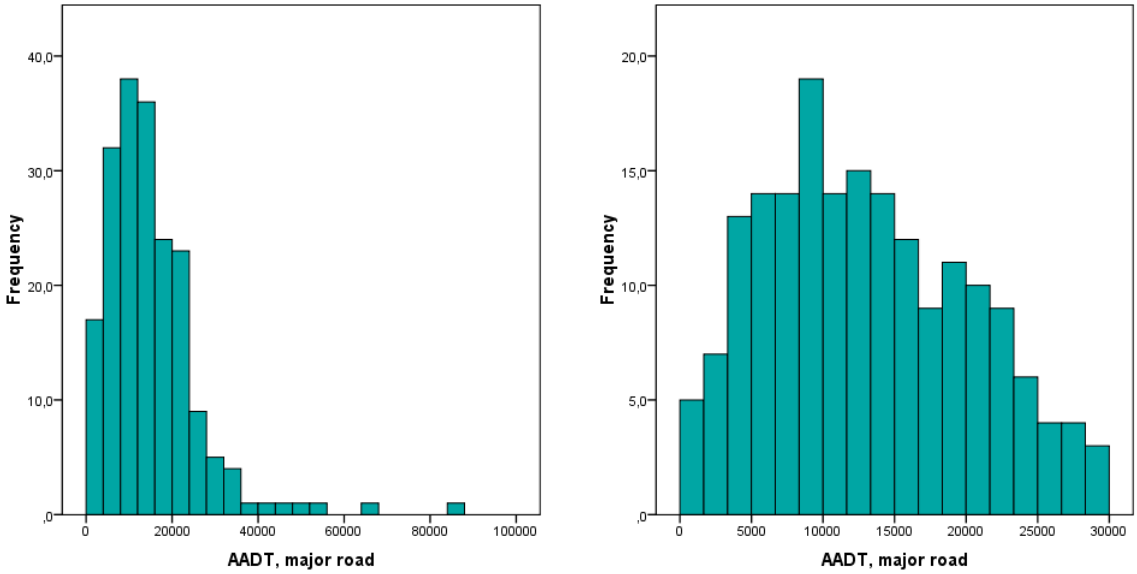


Figure 2 Histogram of AADT on the major roads (All values, values selected for modelling)

Percentage of heavy goods vehicles

Percentage of heavy goods vehicles were taken also from TSK records. It was counted for major and minor roads. The highest value is approximately 35 % for the major road and 75 % in one case for the minor road. Average value is around 7 % for the major roads and 5 % for the minor roads.

Number of legs

Only intersections with three or four legs were used. In total, 119 three leg intersections and 75 four leg intersections were included.

Left turn restrictions

In some intersections, there are legs with one-way traffic or some restricted intersection movements. In 57 cases, there are some restrictions which do not allow turning left in some direction.

Legs with trams and percentage of trams

Two variables for trams were available in the data. The first shows the number of legs of intersections with tram traffic presence. It is smaller or equal to the number of legs of the intersection. Overall, there were 62 intersections with tram traffic. Because of the low number of intersections with trams in three legs, binary information was used for the modeling.
Tram volume was converted to percentage. Maximum value is about 110 % of the AADT on major road and about 193 % of the AADT on minor road. Average value is approximately 4 % for major roads and 3 % for minor roads.

Traffic islands

In some intersections, there are traffic islands that are used to separate the opposite traffic, emphasize the way for turning vehicles or to protect pedestrians crossing the road. Often multiple functions are assigned to them. 45 of intersections have some kind of traffic islands. The number of islands is mostly 0 or 1, higher number of islands is rare, so binary variable was used for modeling.

Priority

Right of way of all intersections is controlled by traffic signs. 46 intersections have major road in a bend. The rest of intersections has the major road direct.

Pedestrian crossing

As the studied intersections are in the city, many of them are equipped with pedestrian crossings. These were divided into groups of crossings across the major and minor road. There are 103 intersections with crossings on the major road and 118 on the minor roads. Test of their significance provided bad results, therefore binary variable (yes or no) was used.

Number of right-turn lanes

Information about number of right turn lanes was available in the data too. According to graph in Figure 3 (frequency of occurrence of intersections with the specific number of right-turn lanes), it is obvious, that intersections with one right-turn lane form 7 % and with two lanes 4 %. As a result, this variable was excluded.

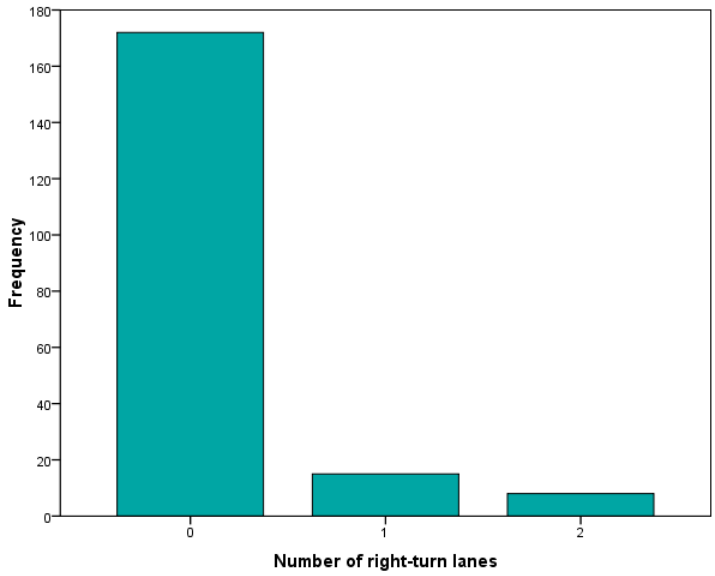


Figure 3 Number of right-turn lanes

Table 1 Overview of variables and their desciptive statistics

Type of variable	**Name of variable**	**Shortcut**	**Data type and unit**	**Descriptive statistics** (min/max/mean/SD or frequencies)
Dependent	Accident data	**ACC**	Count	0.00/29.00/3.22/3.93
Continous	Annual average daily traffic - major road	**AADT_mj**	continuous [vehicles per day]	900/29504/12940/6969
	Annual average daily traffic - minor road	**AADT_mi**	continuous [vehicles per day]	100/14531/3515/3183
	Percentage of HGV	**P_HGV**	continuous	0.008/0.443/0.063/0.047

	Percentage of trams	**P_TRAM**	continuous	0.000/1.005/0.030/0.103
Cathegorical	Number of legs	**LEGS**	nominal	3 : 119, 4 : 75
	Legs with tram traffic	**TRAM**	binary Yes = 1; No = 0	1 : 62, 0 : 132
	Left turn restrictions	**LEFT**	binary No restrictions: 1, With restrictions: 0	1: 137, 0 : 57
	Traffic islands	**ISL**	binary Yes = 1; No = 0	1 : 45, 0 : 149
	Priority	**PRIOR**	binary Bend = 1, Straight = 0	1 : 46, 0 : 148
	Pedestrian crossing	**CROSS**	binary Yes = 1; No = 0	1 : 138, 0 : 56

3) MODEL

Statistical model, based on the described data, was developed in SPSS. Generalized linear, model with negative binomial distribution and logarithmic link function, was used. Accident frequency, from years 2009, 2010 and both of these years, was dependent variable. Process used for development of model was described e. g. by Hauer (2004).

CORRELATION RELATIONSHIPS OF VARIABLES

According to the theory of using statistical modeling for road safety evaluation, all used variables should be independent. Therefore, relationships between variables were also studied and variables with strong correlation were modified.

Obvious correlation is between number of legs of intersection and number of legs from which is possible to turn left. For intersection with three legs, it is usual that from two legs is possible to turn left. For intersections with four legs there are four possible left turns.

There was also found correlation between the number of legs and traffic volume on the major road. This correlation is significant at the 0.05 level and its value is – 0.144, Relationship is logical, more legs would produce more vehicles, so both variables are used in model.

MODEL FORM

The form of the prediction model used in this study was used in other studies as well (E.g. Brüde and Larsson 1993, Miau and Lord 2003, Persaud et al. 2003). Estimated number of accidents is a function of the traffic volume on the major road and minor road and set of risk factors.

General form of the model is:

$$ACC = \beta_0 \cdot AADT_mj^{\beta_1} \cdot AADT_mi^{\beta_2} \cdot e^{\sum \beta_i \cdot x_i}$$

Where *ACC*, *AADT_mj* and *AADT_mi* are described in Table 1 and

β_i … estimated parameters

x_i … other explanatory variables

Model was built by progressive form of building model. The explanatory variable with the highest influence was traffic volume, for that reason it was used first. Than every other variable in specified function form was included in the model and every of these models was evaluated. Three different criteria were checked:

- Decrease of Akaike information criterion (AIC), a measure of the relative goodness of fit of a statistical model
- Decrease of overdispersion parameter
- Shape of cumulative residuals graph created according to Hauer and Bamfo (1997).

Table 2 Final versions of models

Variable		2009			2010			2009-2010		
Shortcut		3 leg intersections								
		Estim. Value	95 % Confid. Interval	Sig.	Estim. Value	95 % Confid. Interval	Sig.	Estim. Value	95 % Confid. Interval	Sig.
Intercept		-7.974	[-10.90;-5.05]	0.000	-8.067	[-10.60;-5.53]	0.000	-7.074	[-9.53;-4.62]	0.000
ln(AADT_mj)		0.670	[0.36;0.98]	0.000	0.480	[0.24;0.72]	0.000	0.594	[0.33;0.86]	0.000
ln(AADT_mi)		0.303	[0.12;0.48]	0.001	0.469	[0.30;0.64]	0.000	0.327	[0.17;0.49]	0.000
P_HGV		-	-	-	0.560	[0.02;0.10]	0.004	4.350	[-0.20;8.89]	0.061
TRAM	Y	0.568	[0.16;0.97]	0.006	0.489	[0.09;0.89]	0.016	0.798	[0.41;1.18]	0.000
	N	0.000			0.000	-	-	0.000	-	-
CROSS	Y	-	-	-	0.392	[0.01;0.78]	0.047	-	-	-
	N	-	-	-	0.000	-	-	-	-	-
		ODP		0.514	ODP		0.431	ODP		0.552

Variable		2009			2010			2009-2010		
Shortcut		4 leg intersections								
		Estim. Value	95 % Confid. Interval	Sig.	Estim. Value	95 % Confid. Interval	Sig.	Estim. Value	95 % Confid. Interval	Sig.
Intercept		-3.626	[-6.71;-0.55]	0.021	-2.530	[-5.12;0.06]	0.056	-3.887	[-7.13;-0.65]	0.019
ln(AADT_mj)		0.321	[-0.01;0.63]	0.044	0.209	[-0.07;0.49]	0.147	0.410	[0.09;0.73]	0.013
ln(AADT_mi)		0.274	[0.12;0.43]	0.000	0.222	[0.08;0.37]	0.003	0.276	[0.14;0.41]	0.000
P_HGV		-0.078	[-0.14;-0.02]	0.007	-0.061	[-0.12;0.00]	0.059	-7.930	[-13.50;-2.36]	0.005
P_TRAM		-	-	-	-	-	-	0.044	[0.01;0.08]	0.019
TRAM	Y	0.497	[0.08;0.91]	0.019	0.576	[0.14;1.01]	0.009	-	-	-
	N	0.000			0.000	-		-	-	-
LEFT	Y	-	-	-	0.643	[0.21;1.08]	0.004	0.496	[0.10;0.89]	0.013
	N	-	-	-	0.000			-	-	-
PRIOR	B	-	-	-	-0.483	[-1.07;0.10]	0.106	-	-	-
	S	-	-	-	0.000	-	-	-	-	-
		ODP		0.432	ODP		0.345	ODP		0.431

Variable		2009			2010			2009-2010		
Shortcut		all intersections								
		Estim. Value	95 % Confid. Interval	Sig.	Estim. Value	95 % Confid. Interval	Sig.	Estim. Value	95 % Confid. Interval	Sig.
Intercept		-6.241	[-8.37;-4.11]	0.000	-4.925	[-6.86;-2.99]	0.000	-4.832	[-6.62;-3.05]	0.000
ln(AADT_mj)		0.534	[0.31;0.76]	0.000	-0.375	[0.12;0.57]	0.000	0.488	[0.29;0.68]	0.000
ln(AADT_mi)		0.295	[0.18;0.41]	0.000	0.310	[0.19;0.43]	0.000	0.264	[0.16;0.37]	0.000
LEGS	3	-0.391	[-0.68;-0.10]	0.008	-0.404	[-0.69;-0.12]	0.006	-0.382	[-0.65;-0.12]	0.005
	4	0.000	-	-	0.000	-	-	0.000	-	-
TRAM	Y	0.580	[0.28;0.88]	0.000	0.640	[0.34;0.94]	0.000	0.635	[0.37;0.90]	0.000
	N	0.000	-	-	0.000	-	-	0.000	-	
LEFT	Y	-	-	-	0.327	[0.03;0.63]	0.000	-	-	-
	N	-	-	-	0.000	-	-	-	-	-
PRIOR	B	-	-	-	-0.315	[-0.69;0.056]	0.10	-	-	-
	S	-	-	-	0.000	-	-	-	-	-
		ODP		0.526	ODP		0.495	ODP		0.560

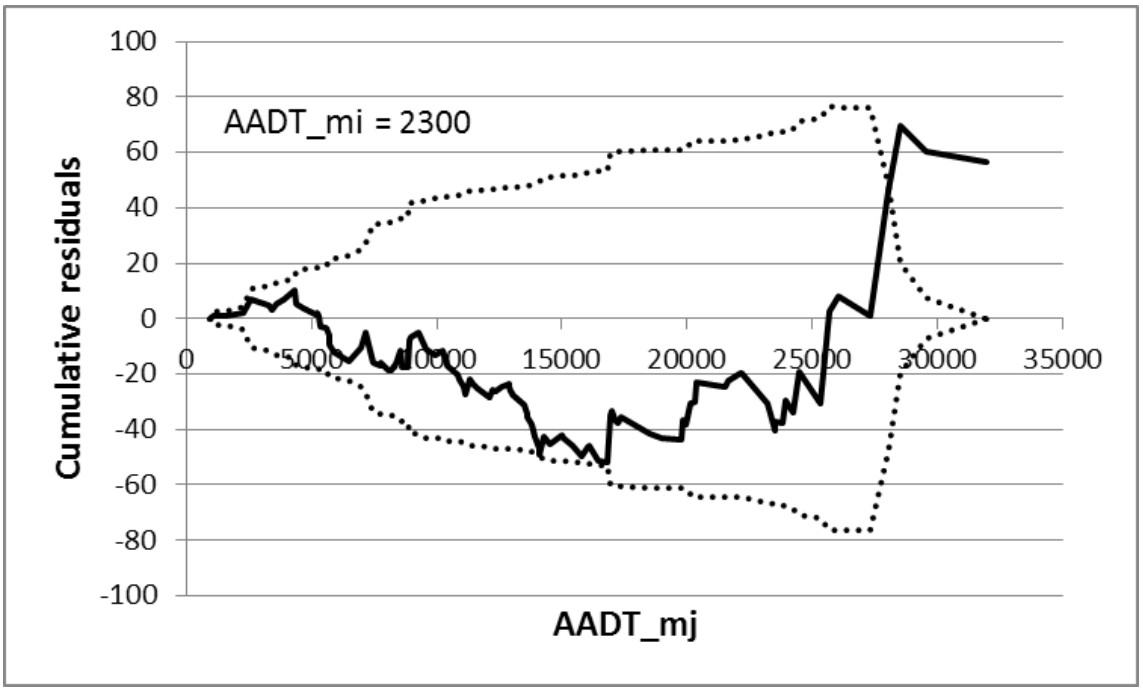


Figure 4 Cumulative residuals for 3 leg intersections for data 2009 - 2010

On graph in Figure 4, the value of AADT_mi is near the median of values of traffic volume on minor roads. The big jump, in the value of cumulative residuals in the interval of volume approximately 27 000 – 29 000, is caused by two intersections where 57 resp. 30 accidents happened. Graph of cumulative residuals of four leg intersections is in the Figure 5.

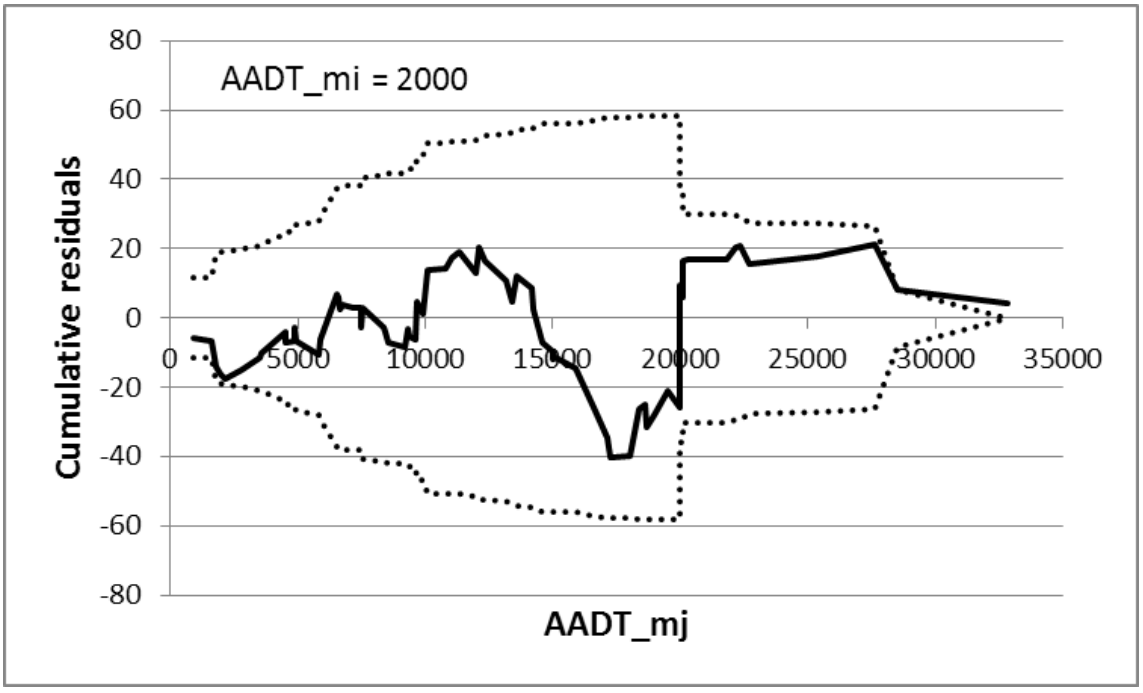


Figure 5 Cumulative residuals for 4 leg intersections for data 2009 - 2010

Final form of the model for three leg intersections is:

$$ACC = e^{-7.074} \cdot AADT_{mj}{}^{0.594} \cdot AADT_{mi}{}^{0.327} \cdot e^{4.350 \cdot p_HGV + 0.798 \cdot TRAM}$$

Final form of the model for four leg intersections is:

$$ACC = e^{-3.887} \cdot AADT_{mj}{}^{0.410} \cdot AADT_{mi}{}^{0.276} \cdot \\ \cdot e^{-7.930 \cdot p_HGV + 0.044 \cdot p_TRAM + 0.496 \cdot LEFT}$$

For both models, binary variables are set 1 for Yes and 0 for No.

4) DISCUSSION

COMPARISON OF MODELS

Models for three leg and four leg intersections were made separately and then third model was made for these types together and the number of legs was one of the variables. In some studies, it is possible to find contention, that separate models for different types of intersections are more reliable (Reurings et al., 2005). In this study the variable LEFT (left turn restrictions) was significant only for four leg intersections. However, it was not significant for model for all intersections together. Also the overdispersion parameter of the models for both types of intersections is lower than the overdispersion parameter of the model for both types together. For that reason, it is possible to submit, that model of each of intersection type is more reliable than one model for all. This conclusion is consistent with cited studies.

Interesting results are in comparison of models in year 2009, 2010 and by using the data for both years. We can find differences in significance of ever of variables. For three leg intersections there is the difference in two variables. For the data of year 2010, there are two more significant variables - percentage of heavy vehicles and the presence of pedestrian crossing. Percentage of heavy vehicles remains significant for the whole two year period, the presence of crossing not. This fact is very interesting. Overdispersion parameter of 2010 dataset is the lowest one. So the model should explain the highest portion of variability in the presence of accidents. Despite this, it is not possible to confirm this statement from the accident records. There are only numbers of accidents in intersections but not their types. Therefore, it is not possible to confirm e.g. that in year 2010 there were higher number of accidents with pedestrians, therefore the presence of crossing should be significant. Very similar situation is for models of both types of intersections. There is one more significant variable for year 2010 - bend priority, however, it is not significant for the two year period. In this case, the idea of too short period is offered. The model shows that the intersection with major road in a bend is more risky than the intersection with major road in straight. This

statement is consistent with general opinion; the intersections with the major road in a bend are more risky. It can be e.g. due to the misunderstanding of drivers about the situation.
Models for four leg intersections shows, why is using of one year period problematic. Behavior of the models is not similar to three leg intersections. The model for year 2010 has difficulties with the reliability. The significance of natural logarithm of the AADT on major road became high, although many other variables are in the limit of 0,100 (10 %). These difficulties did not occur in the model for two year period. Issues identified in the model for year 2010 can be due to deficient number of investigated data, the number of accidents was low too.

EFFECT OF VARIOUS INTERSECTION ARRANGEMENTS

Estimation of safety effect of different types of road safety measures is very complicated. Even though 194 intersections were used for comparison, it does not lend itself to clear interpretation. Many of intersection elements arrangement were not significant and it is likely, that many other parameters have stronger influence on accident rate in the urban area. They can be for example sight conditions, volume of pedestrians and cyclists, presence of advertising or shops and so on.
From the results of this study it is possible to deduce, that the presence of tram traffic increases the risk at intersections. In all cases there was the influence of probability of accidents by tram traffic. In one case the tram traffic volume was significant, in other cases was significant the presence of trams. It does not depend on the numbers of legs of intersection with tram traffic present.
Restrictions of left turn in four legs intersections increase the risk too. It is difficult to say why. Lower number of allowed maneuvers in the intersection should reduce the risk. However the results show the opposite. It could be due to violations of these restrictions, or it is used only in intersections with other risk factors not controlled for in the model; however detailed accident data would be necessary to confirm this idea. Differences can be also found for the type of intersections. Three leg intersections are safer than four legs. This fact is consistent with the number of collision points in each of the intersection type.
This study shows that it is very difficult to find generally valid equation for the specific intersection arrangements in the urban area. There are too many factors, which are not possible to include in the statistical model. Study would have been very complex on large number of intersections with detailed accident data. Causes of accidents would have been normalized; the most significant variables would have been found and used for the modeling.

ACKNOWLEDGMENTS

This article was prepared within the project TA01031303 "Research of effectiveness of suitable junction design parameters using traffic engineering analysis of traffic flow" supported by Technology Agency of the Czech Republic.

REFERENCES

Brüde, U., Larsson, J. (1993). Models for prediction accidents at junctions where pedestrians and cyclist are involved. How well do they fit? *Accident Analysis and Prevention*, **25**, 499-509.

Directive 2008/96/EC of the European Parliament and of the Council of 19 November 2008 on road infrastructure safety management. OJ L 319, 29.11.2008, p. 59–67

European Commission (2001). White paper – European transport policy for 2010: time to decide. COM/2001/0370 final. Available at <http://ec.europa.eu/transport/themes/strategies/doc/2001_white_paper/lb_texte_complet_en.pdf>.

Hauer, E., Bamfo, J. (1997). *Two tools for finding what function links the dependent variable to the explanatory variables.* Proceedings of the ICTCT 1997 Conference, Lund, Sweden.

Hauer, E., Harwood, D. W., Council, F. M., and Griffith, M. S. (2002) Estimating Safety by the Empirical Bayes Method - A Tutorial. *Transportation Research Record*, **1784**, 126-131.

Hauer, E. (2004) Statistical Road Safety Modeling. *Transportation Research Record*, **1897**, 81-87.

Miaou, S-P., Lord, D. (2003). Modeling traffic crash-flow relationships for intersections. Dispersion parameter, functional forma, and Bayes versus empirical Bayes methods. *Transportation Research Record*, **1840**, 31-40.

Persaud, B. et al. (2003). Development of procedure for estimating expected safety effects of contemplated traffic signal installation. *Transportation Research Record*, **1840**, 96-103.

Reurings, M., T. Janssen, R. Eenink, R. Elvik, J. Cardoso, C. Stefan (2005). Accident prediction models and road safety impact assessment: a state-of-the-art. RIPCORD-ISEREST project deliverable D2.1 available at <http://ripcord.bast.de/pdf/RI-SWOV-WP2-R1-State_of_the_Art.pdf>.

Striegler, R., Valentová, V., Pokorný, P., Ambros J., Šenk, P., Janoška, Z. (2012) *Identifikace kritických míst na PK v extravilánu*. CDV, Brno.