ISSN 2669-2023 (print) / ISSN 2538-8959 (on-line) / DOI 10.3849/cndcgs.2024.472

Analysis of Traffic Accidents with the Deployment of the Fire Rescue Service in the Regions of the Czech Republic

Petra KOLISKOVA[1]*, Jiri NEUBAUER[2]

[1]*University of Defence, Kounicova 65, 66210 Brno, Czech Republic*
[2]*Department of Quantitative Methods, Faculty of Military Leadership, University of Defence, Kounicova 65, 66210 Brno, Czech Republic*

Correspondence: *petra.koliskova@unob.cz

Abstract

Traffic accidents are an ongoing safety issue. The aim of this paper is to use a statistical apparatus to describe and model the random variable that is the daily number of accidents. The data on traffic accidents come from the database of the Fire Rescue Service of the Czech Republic from 2012–2021 and are described by time and location data. Analysis of variance for a random variable with a Poisson probability distribution was used in the modelling. The resulting analyses describe the evolution of the daily number of accidents with the deployment of firefighters in each region of the Czech Republic. The work that deals with the modelling of the number of traffic accidents, as a random variable with a Poisson probability distribution, using GLM methods is unique in the Czech Republic.

KEY WORDS: *Traffic accident, fire rescue service, generalised linear models, Poisson distribution.*

Citation: Koliskova, P., Neubauer, J. (2024). Analysis of Traffic Accidents with the Deployment of the Fire Rescue Service in the Regions of the Czech Republic. In Proceedings of the Challenges to National Defence in Contemporary Geopolitical Situation, Brno, Czech Republic, 11–13 September 2024. ISSN 2538-8959, DOI 10.3849/cndcgs.2024.472.

1. Introduction

Traffic accidents are an ongoing safety issue. On the one hand, today's cars are equipped with all sorts of technology and manufactured with much more regard for safety than in the past, but on the other hand, cars are becoming more affordable, traffic is getting heavier and drivers are often encouraged to use high performance cars, which in turn has a negative effect on accident rates. The risk of a traffic accident is one of the most serious incidents investigated by the Integrated Rescue System. In traffic accidents such as those involving extrication [14], the need to deal with leaking fluids, the need to carry out fire-fighting measures on vehicles, or accidents in which it is necessary to restore traffic flow on the motorway, the intervention of fire brigades is required [9].

Many articles from our country and abroad are devoted to the topic of traffic accidents. The spatial representation to identify concentrations of motorcycle accident key hotspots is dealt with in [10]. They use the tool KDE+ (an extension of the kernel density estimation method), which allows to find the most dangerous sections of the Czech road network. Based on traffic accidents in the Czech Republic in 2016–2020, it can be said that traffic accidents are strongly seasonal in terms of time of day and year, and depend on population and traffic density, as well as weather and related conditions.

The road safety strategies for the Czech Republic for the period 2021–2030, presented in [8], are intended to reduce the negative impacts of traffic accidents. The aim of the convention, which is in line with the intentions of the European Union and the United Nations, is to reduce the number of deaths and injuries in road accidents by half. The strategy is a tool to achieve this objective. The key is to shift responsibility from national to local level, allowing concrete actions to be taken at the locations identified by the KDE+ method. Actions include activities in areas such as police surveillance of speed limits, traffic education, awareness of the impact of addictive substances on driving ability, mobile application for central reporting of defects on roads, research on the impact of new modes of transport and alternative vehicle propulsion and automation in relation to road safety, etc.

The topic of firefighting is currently being researched by experts at the University of Bourgogne Franche-Comté in France. The authors draw on data covering firefighting interventions from 2012-2017 in the Doubs department in France. They supplemented the data from the fire department database to 747 statistical features [6], such as meteorological data,

© 2024 The Authors.
Peer-review under responsibility of General Jonas Žemaitis Miltitary Academy of Lithuania and University of Defence, Czech Republic

astrological data, calendar data, and traffic events. In paper [2], the long short-term memory (LSTM) method from the field of artificial neural networks is mentioned. The work is developed in paper [3] and extended by comparing it with another method, extreme gradient boosting (XGBoost). The aim was to make predictions for each year based on data from other years. Predictions for years that were accompanied by natural disasters, and thus unexpected increased rescue activity, were less accurate. In terms of comparing the methods, the results proved almost comparable, with the XGBoost technique better at detecting extreme events during natural disasters. However, the authors also concluded that none of the neural network designs used were ideal due to the limitations of the number of neurons and hidden layers, and thus another approach should continue to be sought to create the best-fitting neural network to describe firefighter interventions.

A comparison of generalised linear model (GLM), i.e., the classical approach using statistics, and artificial neural network (ANN), i.e., the use of machine learning and intelligence, can also be found in [13]. The data of traffic accidents from 2005-2019 that occurred in the Province of Erzurum are considered. The paper describes the geographical characteristics of the study area, including the infrastructure, from which eight independent variables rise that affect the dependent variable: the number of traffic accidents. The calculations showed that ANN gives better results compared to GLM - high correlation coefficient and R-squared values as well as low MSE and RMSE of the tested network, thus confirming its superiority. The graphical representation showed that the ANN model is closer to real traffic accident data than GLM. The aim of this paper was to predict the risk factors and accident frequency. In conclusion, the proposed ANN model attained much better fitting and forecasting functionality compared to the GLM.

Two approaches to modelling the occurrence of traffic accidents are presented in [1]. First, the classical approach of generalised linear regression modelling was chosen for data with distributions such as Poisson, Negative Binomial (NB), Zero Inflated Poisson (ZIP), and Zero Inflated Negative Binomial (ZINB). The second approach, newly introduced, involves normalizing the data and using a linear regression model. This alternative approach was validated on crash data from 186 access roads in the state of Virginia, which was found to be negatively binomial distributed.

The use of a generalised Poisson linear model to evaluate fatal crashes that occurred in Romania between 2008 and 2012 is described in [12]. Factors characterising the crashes were date and time, infrastructure information, cause of crash, safety precautions, and visibility. The paper compares the Poisson and quasi-Poisson approaches of the distribution of the explained variable. As a result, it is concluded that the generalised Poisson linear model allows to select significant risk factors, which in this case turned out to be, among others, the side impact on curved roads.

Among the factors influencing the occurrence of a traffic accident is undoubtedly the influence of weather. The paper [7] analysing traffic accidents in the Chinese city of Shantou examines the influence of meteorological parameters on accident rates by time series methods, correlation analysis and multiple linear regression analysis. The models showed that road traffic injuries are positively correlated with temperature and sunshine duration, while negatively correlated with wind speed.

In the past ten years, a total of 198,773 accidents involving the deployment of fire rescue units occurred in 14 regions of the Czech Republic. The records of accidents in the years 2012-2021 were taken from the database of the Fire Rescue Service of the Czech Republic, where they are characterised by geographical position and time of occurrence. The aim of this work is to use statistical apparatus to describe a discrete random variable, which is the daily number of accidents, and to model it using a generalised linear model (GLM).

First, the probability distribution of the explained random variable was tested. Based on the results of these tests, the analysis of variance approach for statistical analysis with a count variable, following a Poisson distribution, was chosen for modelling. Multifactor analyses without and with interactions were used to describe the daily number of accidents. In this paper, we will consider the following factors: day of the week, month, and region. The results of the estimated models were compared with the statistical characteristics of the dataset. Additionally, it is important to include an interaction term in multifactor analyses.

2. The Mathematical Background

The number of traffic accidents has been successfully described by a random variable with a Poisson, Poisson-gamma, or zero-inflated probability distribution [11]. The Shapiro, Lilliefors, or Anderson-Darling tests can be used to verify the normal probability distribution of a random variable. To verify the Poisson probability distribution, the Cramer-von Mises or Anderson-Darling goodness-of-fit tests can be used. If the p-value of these tests is less than the chosen significance level α, we reject the null hypothesis of a Poisson distribution of the explained variable. Agreement with the assumed distribution can also be verified graphically, e.g., by plotting the relative frequencies of the observed variable or by comparing the empirical and theoretical distribution functions.

The main idea of this analysis is to decide on the dependence of the observed random variable on selected factors. The Y_i values of the explained random variable are sorted into groups according to the variations of the specified factor, and if the Y_i values are shown to differ between these groups, the dependence of the Y variable on the given factor is thereby demonstrated. In the case of a normal distribution of the random variable Y, the dependence on the factors is determined by analysis of variance (ANOVA). Furthermore, the use of multiple comparison tests makes it possible to decide which groups resulting from factor sorting are different. For data that do not follow a normal distribution, the non-parametric Kruskal-Wallis test is used [4]. If the p-values of the Kruskal-Wallis test come out smaller than the chosen significance level, it means that the explained random variable depends on the factor. For the purpose of multiple comparisons, for example, Nemenyi's all-pairs rank comparison test can be used.

For variables that do not meet the conditions for ANOVA, it is also possible to use generalised linear modelling GLM [5]. A random variable Y has a Poisson probability distribution if the probability function can be expressed in the form

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda}, \quad (1)$$

where parameter λ represents the average number of events per time unit and is both the mean and variance of Y. In this case, the natural logarithm is chosen as the link function of the generalised linear model. For the case of a two-factor analysis of variance with interactions, we can write for the means λ_{ij} ($i = 1,2, \dots, I$, and $j = 1,2, \dots, J$,)

$$\log \lambda_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}, \quad (2)$$

where α_i is the effect of the first factor, β_j is the effect for the second factor and γ_{ij} is the interaction effect. From the output of the computational process, it is possible to see how much each factor contributes to the estimate of the mean and how statistically significant their contribution is. The multivariate analysis can be performed in two forms, either without or with factor interactions. In the analysis with interactions, not only the effect of the individual factors on the explained variable is taken into account, but also the effect of the interaction between the different types of factors.

For both models, the percentage of explained variance can be determined by the expression

$$1 - \frac{D}{D_{null}}, \quad (3)$$

where null deviance D_{null} corresponds to a model containing only the intercept (the worst fit), and the residual deviation D is indicative of a model including the independent variables. The higher the D, the larger the difference between the estimated and observed values and thus the model is less able to describe the observed dependence. The percentage of explained deviance indicates how accurate the model is - it takes values from 0 to 1, the closer to 1, the better the model describes reality.

The calculations were performed in R, version 4.3.1.

3. Investigation Results

This article deals with the analysis of traffic accidents with the deployment of firefighters in the Czech Republic in the years 2012-2021. The initial dataset comes from the database of the Fire Rescue Service and includes records of 198,773 accidents that occurred in the Czech Republic from January 1st, 2012 to December 31th, 2021. The colour differentiation of the regions of the Czech Republic according to the number of such accidents can be seen in Fig. 1.

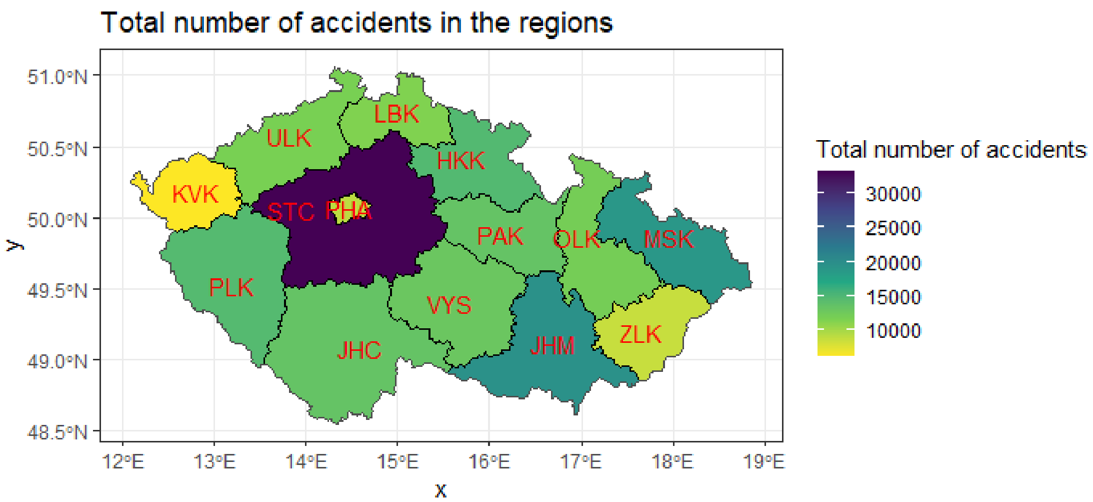

Fig. 1. Total number of accidents with the deployment of firefighters in the regions of the Czech Republic in the years 2012–2021.

Each accident in the database was described with time and location data. The explained coincidence variable is the daily number of traffic accidents with the deployment of firefighters. The observed factors are the *day* and *month* when the accident occurred and the *region* of the Czech Republic in which the accident occurred. The statistical characteristics of accidents in the individual regions of the Czech Republic for the period under study are presented in Table 1 and graphically represented by boxplots in Fig. 2. From Table 1 it can be seen that the region with the highest average daily number of traffic accidents involving deployment of firefighters is the Central Bohemia Region (STC). In contrast, the lowest average number of accidents per day occurred in the Karlovy Vary Region (KVK). In addition to the *mean* of daily number of accidents, the table shows the *minimum* and *maximum* number of accidents per day in a given region, the *median*, and the *standard deviation* of the region's average from the overall daily average. The number of all observations *n*, i.e., all days in the period of interest, is 3653, the same for all regions. The average comes out to 3.89 accidents per day per region.

Table 1.

Characteristics of the daily number of accidents for region.

Region	Mean	Median	Standard deviation	Min	Max
PHA	2.46	2	1.77	0	13
STC	9.08	9	4.61	0	43
JHC	3.65	3	2.33	0	19
PLK	3.98	4	2.56	0	22
KVK	1.66	1	1.48	0	12
ULK	3.19	3	2.08	0	16
LBK	3.10	3	2.24	0	20
HKK	4.02	4	2.56	0	26
PAK	3.65	3	2.82	0	64
VYS	3.53	3	2.62	0	48
JHM	5.34	3	3.28	0	40
OLK	3.28	3	2.28	0	36
MSK	5.15	5	3.00	0	42
ZLK	2.33	2	1.74	0	12

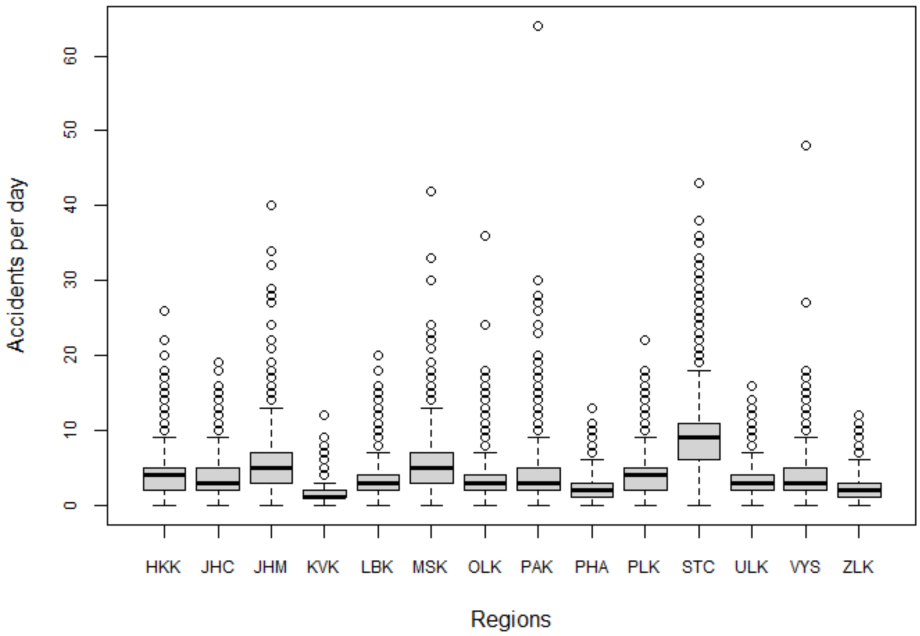

Fig. 2. Boxplots of the daily number of accidents in regions.

Outliers have not been removed from the data as they could be an important indicator of problem locations and periods when extreme numbers of accidents occur. The daily number of crashes in regions is an absolute frequency that has not been converted to region size or population. It is clear that many conditions have a significant effect on the accident rate in regions, e.g., region size, population density, length of the road network, presence of highway segments, etc., which may be the subject of further analysis.

During the initial stages of model development, tests were conducted to evaluate the normal distribution, but the results indicated that this probability distribution would not be appropriate. Therefore, goodness-of-fit tests were performed with the Poisson distribution, namely Cramer-von Mises or Anderson-Darling. Verification was also performed graphically using the values of the estimated likelihood function, or by plotting the relative frequencies for data sorted by region, month, and day of the week. As an example, the comparison plots for January Mondays in the Hradec Kralove region are presented in Fig. 3.

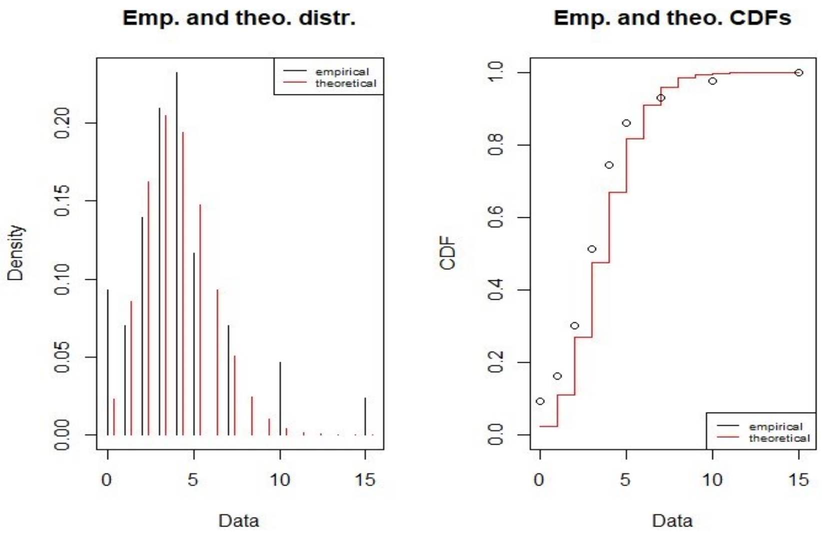

Fig. 3. Probability function and cumulative distribution function for the theoretical Poisson distribution function with estimated parameters and empirical data for accidents in the Hradec Kralove region that occurred on Mondays in January 2012–2021. The theoretical function is shown in red, the empirical function in black.

Since the normality assumptions for parametric analysis of variance are not met, one option is to use non-parametric methods, which is the Kruskal-Wallis test. The Kruskal-Wallis tests were demonstrated for all factors and showed that the p-value is much less than 0.05 for each of the *day* of week, *month* and *region* factors (p-value < 0.001), so that the null hypothesis of equality of means between groups of explanatory variables ordered by the *day* of week, *month* and *region* factors can be rejected. Furthermore, multiple comparison tests were demonstrated for regions, showing that all effects of the *region* factor are statistically significant.

Another option chosen was the generalised linear model approach for count data (Poisson distribution) [5]. This approach allows for both one-factor and multifactor analyses. The parameters estimated according to the one-factor model did not yield any surprising results, they just correspond to the statistical characteristics shown in Table 1.

Table 2.
Generalised linear model – parameter estimates of Model 1 (without interactions).

Factor	*Estimate*	*p-value*		*Factor*	*Estimate*	*p-value*	
Intercept	1.2653	<0.001	***	*HKK*	0.1142	<0.001	***
day1	0.0427	<0.001	***	*JHC*	0.0185	<0.001	***
day2	–0.0144	0.009	**	*JHM*	0.3990	0.027	*
day3	–0.0026	0.637		*KVK*	–0.7686	<0.001	***
day4	0.0253	<0.001	***	*LBK*	–0.1469	<0.001	***
day5	0.1562	<0.001	***	*MSK*	0.3617	<0.001	***
day6	–0.0248	<0.001	***	*OLK*	–0.0903	<0.001	***
day7	–0.1823	<0.001	***	*PAK*	0.0187	<0.001	***
month1	–0.0303	<0.001	***	*PHA*	–0.3777	0.026	*
month2	–0.1052	<0.001	***	*PLK*	0.1040	<0.001	***
month3	–0.2325	<0.001	***	*STC*	0.9297	<0.001	***
month4	–0.1733	<0.001	***	*ULK*	–0.1153	<0.001	***
month5	–0.0490	<0.001	***	*VYS*	–0.0159	<0.001	***
month6	0.1255	<0.001	***	*ZLK*	–0.4310	<0.001	***
month7	0.0713	<0.001	***				
month8	0.1076	<0.001	***				
month9	0.1223	<0.001	***				
month10	0.0872	<0.001	***				
month11	–0.0088	0.242					
month12	0.0853	<0.001	***				

The analysis is mainly focused on a multi-factor model. Parameter estimates were computed for the multivariate analysis of the model with *day*, *month* and *region* factors. Table 2 shows the values of the estimates for the model without interactions

$$\text{Model 1: } \log accidents_{ijk} = \beta_0 + region_i + month_j + day_k, \quad (4)$$

where $i = 1,\dots,14$, $j = 1,\dots,12$, $k = 1,\dots,7$. Table 2 shows that the p-values for most of the factor variations indicate high statistical significance. The parameter estimates can be used to calculate an estimate of the average value of the daily number of traffic accidents in a particular *day*, *month* and *region*. For example, effect for Monday is exp(0.0427). Thus, the estimate of the daily number of accidents on a Monday in January in the Hradec Kralove region (HKK) is equal to exp(1.2653+0.0427–0.0303+0.1142) = 4.022.

The average daily numbers of accidents in each day and month were calculated for all regions. The tables with the calculated averages for the region with the lowest average daily number of accidents with fire brigade deployment, i.e. the Karlovy Vary Region (KVK) in Table 3, and for the region with the highest average daily number of accidents, i.e. the Central Bohemia Region (STC) in Table 4, are attached for your reference.

Table 3.

Predicted values of the daily number of accidents, Model 1 (without interactions) – KVK, the region with the lowest average daily number of accidents. The colour scale was chosen from green (= lowest value) to red (= highest value).

Month\Day	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
1	1.66	1.57	1.59	1.64	1.86	1.56	1.33
2	1.54	1.46	1.48	1.52	1.73	1.44	1.23
3	1.36	1.28	1.30	1.34	1.52	1.27	1.09
4	1.44	1.36	1.38	1.42	1.62	1.35	1.15
5	1.63	1.54	1.56	1.60	1.83	1.53	1.30
6	1.94	1.84	1.86	1.91	2.18	1.82	1.55
7	1.84	1.74	1.76	1.81	2.06	1.72	1.47
8	1.91	1.80	1.83	1.88	2.14	1.79	1.53
9	1.94	1.83	1.85	1.90	2.17	1.81	1.55
10	1.87	1.77	1.79	1.84	2.10	1.75	1.49
11	1.70	1.61	1.62	1.67	1.90	1.59	1.36
12	1.87	1.76	1.78	1.84	2.09	1.75	1.49

Table 4.

Predicted values of the daily number of accidents, Model 1 (without interactions) – STC, the region with the highest average daily number of accidents.

Month\Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	9.09	8.59	8.69	8.94	10.18	8.50	7.26
2	8.44	7.97	8.06	8.29	9.45	7.88	6.74
3	7.43	7.01	7.10	7.30	8.32	6.94	5.93
4	7.88	7.44	7.53	7.74	8.83	7.37	6.29
5	8.92	8.43	8.53	8.77	10.00	8.34	7.13
6	10.62	10.03	10.15	10.44	11.90	9.93	8.48
7	10.06	9.51	9.62	9.89	11.27	9.41	8.04
8	10.44	9.86	9.97	10.26	11.69	9.75	8.33
9	10.59	10.00	10.12	10.41	11.86	9.90	8.46
10	10.22	9.66	9.77	10.05	11.45	9.56	8.16
11	9.29	8.77	8.88	9.13	10.40	8.68	7.42
12	10.21	9.64	9.75	10.03	11.43	9.54	8.15

It can be seen that the factor *day* of the week showed that in terms of accidents, Mondays, Thursdays and Fridays are above average, while on Sundays the risk of a traffic accident involving the deployment of the fire rescue service was below average. The development of the risk of traffic accidents during the year, based on the collected data sample, shows that from November

to May, except December, the risk is below average, while from June to October and in December the risk is above average, for both displayed regions. Similar results emerge for the remaining regions of the Czech Republic.
The characteristics of the model without interactions resulting from the use of the GLM method are summarized in Table 5. To test the null hypothesis that the model describes the data well, the *deviance* statistic is used, which has an approximate χ^2 distribution with *Df* degrees of freedom. Here the model is compared to the saturated (maximum) model, which can be seen in the *Resid. Dev* value. The higher the deviance values, the worse the model. Thus, Table 5 shows that the percentage of explained deviance is 32.54%.

Table 5.
ANOVA – Model 1 (without interactions).

	Df	**Deviance**	**Resid. Df**	**Resid. Dev**	**Pr(>Chi)**
null			51141	117557	
region	13	34050	51128	83507	<0.001
day	6	1689	51111	79305	<0.001
month	11	2512	51117	80995	<0.001

Several possible interactions between factors have been studied, of which the *month* and *day* interaction seems to be the most meaningful. We apply the model

$$\text{Model 2: } \log accidents_{ijk} = \beta_0 + region_i + month_j + day_k + month \times day_{jk}, \quad (5)$$

where $i = 1,\ldots,14$, $j = 1,\ldots,12$, $k = 1,\ldots,7$. This model with the interaction is shown in Table 6, a comprehensive table would contain 1176 values. The percentage of deviance explained by this model is 0.330. A comparison of the two models is made in Table 7, showing that the multi-factor model with interactions seems to be a suitable model for analysing the dependence of the daily number of accidents involving the deployment of the fire rescue service according to the *day* of the week, *month* and *region*, as evidenced by the very small p-value.

Table 6.
ANOVA – Model 2 (with interactions)

	Df	**Deviance**	**Resid. Df**	**Resid. Dev**	**Pr(>Chi)**
null			51141	117557	
region	13	34050	51128	83507	<0.001
day	6	1689	51111	79305	<0.001
month	11	2512	51117	80995	<0.001
month×day	66	539	51045	78766	<0.001

Table 7.
ANOVA – comparison of Model 1 and Model 2

	Df	**Deviance**	**Resid. Df**	**Resid. Dev**	**Pr(>Chi)**
1			51111	79305	
2	66	539.1	51045	78766	<0.001

For comparison, estimates of the daily means of traffic accidents were also calculated for the model with interactions, see Table 8 and 9. It can be seen that the model with interactions shows deeper differences between days, while the model without interactions shows more moderate values that are closer to the overall average.

Table 8.
Predicted values of the daily number of accidents, Model 2 (with interactions) – KVK,
the region with the lowest average daily number of accidents.

Month\Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	1.69	1.75	1.69	1.63	1.90	1.42	1.12
2	1.51	1.53	1.73	1.43	1.78	1.28	1.13
3	1.34	1.42	1.31	1.33	1.51	1.22	1.02
4	1.45	1.40	1.37	1.51	1.63	1.29	1.07
5	1.67	1.55	1.52	1.66	1.78	1.49	1.33
6	1.94	1.80	1.72	1.86	2.23	1.91	1.63

7	1.80	1.64	1.76	1.84	2.00	1.82	1.56
8	1.85	1.74	1.79	1.78	2.10	1.88	1.72
9	1.98	1.65	1.76	1.78	2.31	1.88	1.69
10	1.79	1.74	1.82	1.86	2.01	1.81	1.57
11	1.79	1.61	1.59	1.73	1.85	1.58	1.30
12	1.91	1.72	1.75	1.94	2.11	1.78	1.38

Table 9.

Predicted values of the daily number of accidents, Model 2 (with interactions) – STC, the region with the highest average daily number of accidents.

Month\Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	9.25	9.55	9.25	8.89	10.37	7.75	6.13
2	8.23	8.35	9.48	7.84	9.74	6.99	6.18
3	7.33	7.78	7.16	7.26	8.28	6.64	5.60
4	7.90	7.65	7.47	8.23	8.91	7.04	5.86
5	9.10	8.49	8.29	9.07	9.74	8.15	7.26
6	10.62	9.82	9.37	10.18	12.21	10.42	8.90
7	9.82	8.98	9.60	10.03	10.91	9.94	8.55
8	10.13	9.53	9.80	9.71	11.46	10.27	9.41
9	10.82	9.02	9.60	9.74	12.64	10.27	9.23
10	9.77	9.49	9.95	10.18	10.99	9.91	8.60
11	9.77	8.82	8.69	9.47	10.09	8.65	7.08
12	10.45	9.41	9.55	10.58	11.52	9.71	7.55

4. Conclusions

The aim of the study was to verify the dependence of the daily number of traffic accidents with the deployment of firefighters on the selected factors, which were the day of the week, month and region of the Czech Republic, using the methods of generalised linear modelling GLM for count variable with Poisson distribution. The results obtained were compared with the origin dataset and a conclusion was drawn based on this comparison.

The hypothesis of normality of the data was rejected after performing tests of normal distribution. Goodness-of-fit tests with Poisson distribution were performed for data sorted by region, month, day of the week and showed approximately good results. For the Kruskal-Wallis tests a p-value much smaller than 0.05 was obtained, based on these tests it can be concluded that the daily number of accidents depends on the observed factors. Further models showed that all factors examined were statistically significant. Multifactor analyses were performed, additive (without interactions between factors) and multiplicative (with interactions between day and month).

Both multifactor models, with and without interactions, were compared. Predicted values of daily number of accidents in the model without interactions are presented for the regions with the lowest and highest daily average number of accidents. In the case of multifactor models, the addition of an interaction term appears to be significant. The percentage of explained variance for the model without interactions is 32.54% and for the model with interactions is 33.00%, i.e., the model with interactions gives slightly better results for describing the observed variables.

The following conclusions emerged from the analyses. The day of the week factor showed that Mondays, Thursdays and Fridays were above average in terms of accident rates, while Sundays were below average in terms of the daily number of traffic accidents involving the deployment of the fire rescue service. Overall, the number of accidents is above average on weekdays compared to the weekend when accidents occur less frequently. The trend in the daily number of traffic accidents over the year based on the data sample collected shows that the daily number is below average in the winter and spring months, and from June to October the accident rate is above average relative to the overall annual average daily number of traffic accidents involving the deployment of the fire rescue service. December showed high values for the daily number of traffic accidents as well. These results reflect the busyness of traffic and how it changes during the week, throughout the year. Quieter days are days when people do not have to commute, for example for work. Summer is the time of year for travel, and high summer temperatures and long journeys have an adverse effect on drivers' attention spans. In the event of an accident, there is a greater need for fire-fighting measures in hot weather than in winter, for example. The dependence of the daily number of accidents on the region also proved statistical significance. The calculations show the lowest average of accidents per day in the Karlovy Vary Region and the highest in the Central Bohemia Region. The region factor was reflected in the accident rate mainly by the geographical conditions in the regions. Regions, where the winters are colder, show a higher accident rate, e.g. due to black ice.

Many factors play a role in traffic accidents. We have focused on the occurrence of traffic accidents in particular days, months and regions, however, other factors that certainly influence the number of traffic accidents with the deployment of firefighters are road infrastructure, region population, geographical characteristics of the region, but also the influence of weather and other meteorological phenomena, such as those mentioned in studies [6], [7], and impact of natural events, e.g. black ice, which are difficult to predict, because they vary from year to year and their influence on the prediction of the number of traffic accidents is rather unpredictable, see [2], [3].

It is hoped that, as a result of the above-mentioned studies, models of the development of the daily number of traffic accidents with the deployment of firefighters will be able to be useful in the field of traffic safety in the Czech Republic. The study provides a comprehensive view of the relationship between the number of accidents and days of the week, months of the year and regions of the Czech Republic. In the event of accidents, significant losses occur, not only financially, but also, and above all, in terms of damage to health and environmental damage. Reducing the need for deployment of firefighters would lead, among other things, to savings in terms of the equipment used and the costs of subsequent treatments and repairs [8]. Based on the knowledge of the development of the daily number of accidents, it is then possible to propose appropriate solutions to reduce the accident rate.

Acknowledgements

This paper was supported from the research project Conduct of Land Operations, LANDOPS, Ministry of the Defence of the Czech Republic.

References

1. **Arafah, M., Abdel-Salam, A.-S., Guo, F., Medina, A.** Linear regression crash prediction models: Issues and proposed solutions. 2010. 10.1201/9780203881200.ch25.
2. **Cerna Nahuis, S., Guyeux, C., H. Arcolezi, H., Couturier, R., Royer, G., Lotufo, A.** Long Short-Term Memory for Predicting Firemen Interventions, 2019. 1132-1137. 10.1109/CoDIT.2019.882067
3. **Cerna Nahuis, S., Guyeux, C., H. Arcolezi, H., Couturier, R., Royer, G.** A Comparison of LSTM and XGBoost for Predicting Firemen Interventions, 2020. 10.1007/978-3-030-45691-7_39.
4. **Devore, J. L.** Probability and Statistics for Engineering and the Sciences. 8th ed. Boston: Brooks/Cole, 2012.
5. **Dobson, A.** An Introduction to Generalised Linear Models. London: Chapman & Hall, 2008.
6. **Guyeux, C. et al.** Firemen Prediction by Using Neural Networks: A Real Case Study. In: Bi. Y., Bhatia. R., Kapoor. S. (eds) Intelligent Systems and Applications. IntelliSys 2019. Advances in Intelligent Systems and Computing. vol 1037. Springer. Cham, 2020. https://doi.org/10.1007/978-3-030-29516-5_42
7. **Gao, J., Chen, X., Woodward, A. et al.** The association between meteorological factors and road traffic injuries: a case analysis from Shantou city. China. Sci Rep 6. 37300, 2016. https://doi.org/10.1038/srep37300
8. **Kadula, L., Valentova, V.** Czech Road Traffic Safety Strategy 2021–2030. XXVII World Road Congress - Prague 2023, 2023. 10.13140/RG.2.2.25438.20804.
9. **Kocinova, M., Nedelnikova, H.** Statistical Yearbook 2021. Ministry of the Interior, Fire Rescue Service of the Czech Republic, 2021.
10. **Kraft, S., Marada, M., Petricek, J., Blazek, V., Mrkvicka, T.** Identification of Motorcycle Accidents Hotspots in the Czech Republic and Their Conditional Factors: The Use of KDE+ and Two-step Cluster Analysis. The Geographical Journal. 2022. 188. 10.1111/geoj.12446.
11. **Lord, D., Washington, S., Ivan, J.** Poisson, Poisson-gamma and zero inflated regression models of motor vehicle crashes: Balancing statistical fit and theory. Accident; analysis and prevention, 37: 35-46. 2005. 10.1016/j.aap.2004.02.004.
12. **Pop, D.** Generalised Poisson Linear Model for Fatal Crashes Analysis. In: Romanian Journal of Automotive Engineering, vol. 24-4, p. 131-134. 2018. ISSN 2457 – 5275 (Online, English)
13. **Sahraei, M. A. et al.** Forecasting the Accident Frequency and Risk Factors: A Case Study of Erzurum, Turkey, Technical Gazette 29, 1(2022), 190-199. https://doi.org/10.17559/TV-20200620164552
14. **Wargclou, D.** Extrication from Cars during Road Traffic Accidents. MSB, 2011. ISBN: 978-91-7383-165-9.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of CNDCGS 2024 and/or the editor(s). CNDCGS 2024 and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.