Exploration of real-time crash likelihood of Powered-Two Wheelers in Greece

Theofilatos Athanasios[1]*, Yannis George[1], Kopelias Pantelis[2], Papadimitriou Fanis[3]

[1]National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Heroon Polytechniou str., Athens, GR15773, Greece
[2]University of Thessaly, Department of Civil Engineering, Pedion Areos, Volos, GR38334, Greece
[3]Attica Tollway Operations Authority – Attikes Diadromes S.A., 41.9 km Attiki Odos Motorway, Paiania, GR19002, Greece

*Corresponding author; e-mail address: atheofil@central.ntua.gr

Abstract

The incorporation of real-time traffic and weather data has proven to be a very fruitful approach when analysing crash likelihood. A major limitation is that there is no specific focus on vulnerable road users such as Powered-Two-Wheelers (PTWs). This paper aims to analyse PTW crash likelihood in the motorway of Attica Tollway ("Attiki Odos") by using real-time traffic and weather data and applying Bayesian Logistic Regression. The results of the paper attempt to contribute to the understanding of accident probability and severity on motorways, by having a special consideration of PTWs for one of the first times for safety evaluation of motorways.

Keywords

Road safety; real-time data; crash likelihood; powered-two-wheelers

1. INTRODUCTION

The European Union (EU) has made substantial progress in improving road safety and reducing traffic fatalities. In the decade leading to 2010 the number of fatalities decreased by 45% and the total injured casualties by 30%. Nevertheless, in 2010, 31,000 fatalities still occurred on EU roads.

Consequently, Understanding the numerous factors that affect road crashes has attracted a great attention in literature and recent studies concentrate on examining the combined effects of traffic and weather on road safety to assist in developing real-time traffic management strategies. Although much progress has been carried out, the vast majority of relevant literature has a focus on US, China and Japan freeways and expressways (Abdel-Aty and Pande, 2005; Abdel-Aty et al., 2007; Ahmed and Abdel-Aty, 2012; Kockelman and Ma, 2007; Yu and Abdel-Aty, 2013; Xu et et., 2013a and 2013b; Hossain and Muromachi 2012 and 2013; Yu et al., 2015; Peng et al., 2017).

It is obvious that European motorways are less explored; let alone less developed countries. Another limitation is the fact, that there is no specific focus on vulnerable road users such as Powered-Two-Wheelers (PTWs). Due to the low weight and the high manoeuvre capabilities of PTWs, and considering the lack of protection of riders, it is of high importance to understand the influence of traffic conditions on PTWs safety.

This paper aims to analyse crash probability of Powered-Two-Wheelers (PTW crash involvement) in the motorway of Attica Tollway ("Attiki Odos") by using real-time traffic and weather data. Crash probability of PTWs is the likelihood that a PTW is involved in a crash given that this crash has already occurred. In order to achieve this aim, Bayesian Logistic models have been applied.

2. DATA

In this paper, the required accident, traffic and weather data were extracted from Attica Tollway Only basic motorway segments (BFS) were considered and not ramp areas. The raw motorway dataset, includes 387 cases, from 2006 to 2011. In order to explore the probability of PTW crash involvement, a subset had to be created. More specifically, all crash cases had to be defined (and not the injured persons). Therefore, each row of this subset corresponds to a crash, resulting in 285 crash cases, where a new binary variable was defined, namely "PTW crash involvement". This variable takes two possible values 1 if a PTW was involved in a crash and 0 if no PTWs were involved in that crash. It is very interesting that PTWs are involved in almost half of the total crash cases (49.5%).

In the present study, crash and real-time traffic data were collected from Attica Tollway ("Attiki Odos"), an urban motorway located in Greater Athens area, Greece, which connects Athens-Lamia National Road with Athens-Corinth National Road, by-passing the centre of Athens. Attica Tollway is one of the largest ring roads in Europe. It is a modern motorway, with a length of 65.2 Km and two directionally separated carriageways, each consisting in three lanes and an emergency lane. Three datasets were used in this analysis: one dataset with crash data, one with traffic data and one with weather data. The required crash data for Attica Tollway were extracted from the Greek crash database SANTRA provided by the Department of Transportation Planning and Engineering of the National Technical University of Athens.

Real-time traffic data for the Attica Tollway were collected from the Traffic Management and Motorway Maintenance. Inductive loops (sensors), placed every 500 meters inside the asphalt pavement of the open sections of the motorway and every 60 meters inside tunnels, are used to provide information regarding the volume, speed and density of traffic. Traffic flow, occupancy, speed and truck proportion were considered and were measure in 5-min intervals and each crash was assigned to the closest upstream loop detector. Real-time weather data were extracted from the website of the Hydrological Observatory of Athens (HOA), whose address is www.hoa.ntua.gr. The site provides hydrological information and is operated by the National Technical University of Athens. It consists of more than 10 stations located in the greater Athens area, measuring various environmental parameters. In our study, each crash was assigned to the closest meteorological station and rainfall, relative humidity and wind speed were utilized.

3. METHODOLOGY

In this paper several Bayesian logistic regression models were developed to estimate the effect of traffic states on accident severity and probability with focus on PTWs. The classical statistical approach (also called frequentist approach) is different than the Bayesian approach. The general philosophy behind Bayesian approach, is that the prior distributions for each parameter are defined and then the data are used to update beliefs about the behaviour of parameters. Moreover, the updated probability of the parameters are used and the posterior credible intervals are produced. The correct interpretation is that a parameter of interest lies within the credible interval with 95% probability. In that context, instead of a t-test, each parameter is statistically significant if the 95% credible interval (2.5%-97.5%) of the beta coefficient does not contain zero (Lunn et al., 2012). As stated by some studies (Ahmed et al., 2012), the Bayesian inference can effectively treat over fitting problems.

Bayesian inference for logistic regression follows the usual procedure for all Bayesian analysis. More specifically, a prior distribution for all unknown parameters has to be formed, then the likelihood function of the data has to be defined and lastly, the Bayes theorem has to be applied so as to find the posterior distribution of all parameters.

The likelihood function for Bayesian logistic regression is the same as in the frequentist inference. More specifically,

$$likelihood_i = \pi(x_i)^{y_i}(1-\pi(x_i))^{(1-y_i)} \quad \text{(Eq. 1)}$$

where $\pi(x_i)$ is the probability of the event for the i^{th}subject which has covariate vector x_i. The y_i is the response variable which has the outcomes y=1 (occurrence of event) or y=0 (absence of event). The logistic regression equation is:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \quad \text{(Eq. 2)}$$

where β_0 is the intercept, β_i is a coefficient for the explanatory variable x_i. In addition, similarly to the frequentist approach, taking the exp(β) provides the odds ratio for one unit change of that parameter.

Any prior distribution can be used for unknown parameters, however, it is usually preferable to use the so-called "vague" or "non-informative" priors if little is known about the coefficient values (Lunn et al., 2012). A non-informative prior could be for example a normal distribution with zero mean and very large variance, for example its form could be: $\beta_j \sim normal(0, 100^2)$. Another popular non-informative prior could be for example a uniform distribution with large boundaries a and b, e.g. $\beta_i \sim uniform(-100, 100)$.

It is noted that in the vast majority of cases when choosing a normal distribution as a prior distribution, the precision is considered. The precision is defined as $\frac{1}{\sigma^2}$, where σ^2 is the variance. Therefore, the distribution of the aforementioned example is transformed to $\beta_j \sim normal(0, 0.0001)$. Lastly, the posterior distribution is derived if the prior distribution over all parameters is multiplied by the full likelihood function. Thus,

$$prior \times likelihood = posterior \qquad \text{(Eq. 3)}$$

4. RESULTS

The relationship between traffic and weather parameters and PTW accident probability, was examined through the application of Bayesian logit models. The followed methodological approach was the same as in previous models. The priors for the constant and for the candidate independent variables were all "vague" (non-informative), assuming to follow a normal distribution with zero mean and very low precision. The prior for the constant was alpha~dnorm(0, 0.0001). All candidate independent variables were following a non-informative normal distribution, e.g. beta~dnorm(0, 0.0001). The first 5,000 samples were discarded as adaptation and burn-in. Three chains and 20,000 more samples were used to ensure convergence. The Monte Carlo (MC) errors (i.e. the Monte Carlo standard error of the mean values) were monitored. According to Spiegelhalter et al. (2003), MC errors less than 0.05 indicate that convergence may have been achieved. In the model all MC errors were very low (less than 0.005) indicating convergence. Figure 1 summarizes the findings of the Bayesian logit model for PTW accident probability, and provides the posterior mean, the standard deviation and the 95% credible interval CI (2.5%-97.5%) and the odds ratios (OR). Only statistical significant parameters are illustrated on the table.

Model1	Parameters Estimates			Credible Intervals		Model2	Parameters Estimates			Credible Intervals	
	Mean	St.Deviation	Odds Ratio	2.50%	97.50%		Mean	St.Deviation	Odds Ratio	2.50%	97.50%
constant	-0.8357	0.2355	-	-1.2970	-0.3816	constant	-1.6950	0.3958	-	-2.5090	-0.9607
Q_avg_30m_up	0.0130	0.0033	1.0131	0.0068	0.0194	Q_avg_30m_up	0.0435	0.0114	1.0445	0.0223	0.0666
						Q_avg_30m_up^2	-0.0002	0.0001	0.9998	-0.0003	-0.0001
DIC	376.926					DIC	367.84				

Figure 1: Significant parameters estimates, credible intervals and odds ratios for PTW accident probability model.

Only the 30-min average traffic flow was found to be statistically significant and is interesting that a quadratic relationship was revealed as well. More specifically, two models were developed, one linear and one non-linear. The equations are provided below:

$$\text{U1} = -0.8357 + 0.013 * \text{Q\_avg\_30m\_up} \qquad \text{(Eq. 4)}$$

$$\text{U2} = -1.695 + 0.0435 * \text{Q\_avg\_30m\_up} - 0.0002 * \text{Q\_avg\_30m\_up}^2 \qquad \text{(Eq. 5)}$$

In the first model the average flow has a positive relationship with PTW accident involvement, suggesting that as traffic flow increases, PTWs are more likely to be involved in accidents. However, in the second model a quadratic term of the flow was found to be significant, implying a non-linear relationship between flow and PTW accident probability. The DIC of the non-linear model is lower, suggesting that this model is preferable over the linear model. Therefore, a quadratic relationship between PTW accident probability and average flow is more likely to exist. The next figures illustrate a graphical representation of the relationship between average flow and the utility function as well as the probability of PTW accident involvement. More specifically, figures 2 and 3 regard the linear model while figures 4 and 5 the non-linear model.

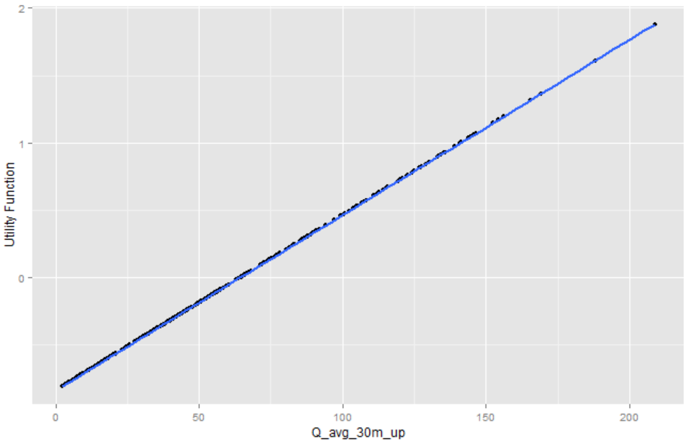

Figure 2: Diagram of the relationship between the average flow upstream and the utility function of the linear PTW crash likelihood model.

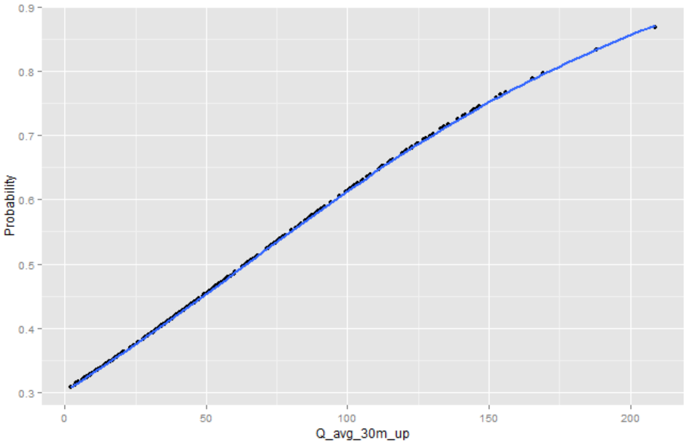

Figure 3: Diagram of the relationship between the average flow upstream and the probability of PTW crash likelihood (linear model).

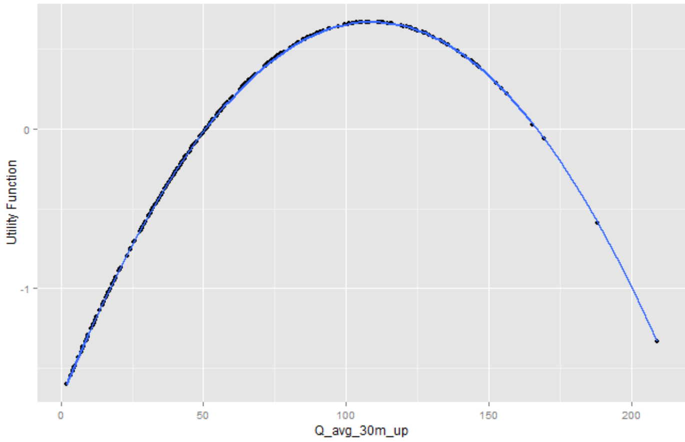

Figure 4: Diagram of the relationship between the average flow upstream and the utility function of the non-linear PTW crash likelihood model.

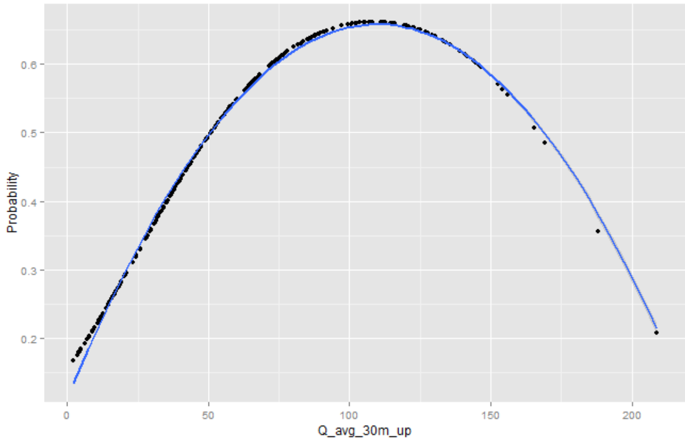

Figure 5: Diagram of the relationship between the average flow upstream and the probability of PTW crash likelihood (non-linear model).

Figure 5 implies that an inverse U-shape relationship exists, meaning that as average flow increases the probability of PTW accident involvement increases until it reaches a maximum and then it starts to decrease.

5. CONCLUSIONS

The aim of the present paper was to investigate road safety in motorways by utilizing high resolution (real-time) traffic and weather data for Attica Tollway ("Attiki Odos") in Greater Athens Area, Greece. The dataset included 285 crash cases (crashes), from 2006 to 2011, in order to explore the probability of PTW crash involvement. Therefore, each row of this subset corresponds to a crash, resulting in 285 crash cases, where a new binary variable was defined, namely "PTW crash involvement".

PTW crash involvement (or PTW crash likelihood) was explored through Bayesian logistic regression models. It was found the only statistically significant variable is the average 30-min flow upstream of the accident location. Two models were developed: one linear and one non-linear. The fit of the non-linear was better indicating that a quadratic relationship exists, namely an inverse U-shape. It is interesting, that weather parameters were not found to significantly affect injury severity of occupants in motorways. The insignificance of weather parameters in the motorway, may be attributed to the fact that weather parameters may not be linearly related with road safety indicators such as PTW probability. It is expected that complex non-linear relations may exist and need further investigation.

Overall, this paper contributes on the current knowledge, by having a specific consideration of PTW safety in motorways and also by developing models combined with real-time traffic and weather data.

Acknowledgements

This research is implemented through IKY scholarships programme and co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the action entitled ”Reinforcement of Postdoctoral Researchers”, in the framework of the Operational Programme ”Human Resources Development Program, Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF) 2014 – 2020. The authors would like to explicitly thank George Yannis, Professor of the National Technical University of Athens, for his guidance and advice for improving the work carried out in this research.

6. REFERENCES

Abdel-Aty M., Pande A. (2005) “Identifying crash propensity using specific traffic speed conditions”, Journal of Safety Research 36, pp. 97–108.

Abdel-Aty M., Pande A., Lee C., Gayah V., Dos Santos, C. (2007) “Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways”, Journal of Intelligent Transportation Systems: Technology, Planning, and Operations 11 (3), pp. 107–120.

Ahmed, M., Abdel-Aty, M., Yu, R. (2012) “Assessment of interaction of crash occurrence, mountainous freeway geometry, real-time weather, and traffic data”, Transportation Research Record, 2280, pp. 51–59.

Hossain, M., Muromachi, Y. (2012) “A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways”. Accident Analysis and Prevention 45, pp. 373–381.

Hossain, M., Muromachi, Y. (2013) “Understanding crash mechanism on urban expressways using high-resolution traffic data”, Accident Analysis and Prevention 57, pp. 17–29.

Kockelman, K., Ma, J. (2007) “Freeway speeds and speed variations preceding crashes, within and across lanes”, Journal of Transportation Research Forum 46 (1), pp. 43–61.

Lunn, D.J., Jackson, C., Best, N., Thomas, A., Spiegelhalter, D. (2012) “The BUGS Book: A practical introduction to Bayesian Analysis”. 1st ed., Boca Raton, FL., Chapman and Hall/CRC.

Peng, Y., Abdel-Aty, M., Shi, Q., Yu, R. (2017) “Assessing the impact of reduced visibility on traffic crash risk using microscopic data and surrogate safety measures”, Transportation Research Part C 74, pp. 295–305.

Spiegelhalter, D., Best, N., Carlin, B., Linde, V. (2003) “Bayesian measures of model complexity and fit (with discussion)”. Journal of the Royal Statistical Society B 64 (4), pp. 583–616.

Xu, C., Tarko, A., Wang, W., Liu, P. (2013b) “Predicting crash likelihood and severity on freeways with real-time loop detector data”, Accident Analysis and Prevention 57, 30–39.

Xu, C., Tarko, A., Wang, W., Liu, P. (2013a) “Identifying crash-prone traffic conditions under different weather on freeways”, Journal of Safety Research 46, pp. 135–144.

Yu, R., Abdel-Aty, M. (2013). “Investigating the different characteristics of weekday and weekend crashes”, Journal of Safety Research 46, pp. 91–97.

Yu, R., Wang, X., Abdel-Aty, M. (2015) “A Hybrid Latent Class Analysis Modeling Approach to Analyze Urban Expressway Crash Risk”, Accident Analysis and Prevention 101, pp. 37-43.