Journal of Safety Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials

Athanasios Theofilatos

National Technical University of Athens, School of Civil Engineering, Dept. of Transportation Planning and Engineering, 5, Iroon Polytechneiou Str., Zografou Campus, Zografou-Athens GR-15773, Greece

A R T I C L E   I N F O

Article history:
Received 15 February 2016
Received in revised form 14 October 2016
Accepted 16 February 2017
Available online xxxx

Keywords:
Accident likelihood
Accident severity
Real-time data
Urban arterials

A B S T R A C T

Introduction: The effective treatment of road accidents and thus the enhancement of road safety is a major concern to societies due to the losses in human lives and the economic and social costs. The investigation of road accident likelihood and severity by utilizing real-time traffic and weather data has recently received significant attention by researchers. However, collected data mainly stem from freeways and expressways. Consequently, the aim of the present paper is to add to the current knowledge by investigating accident likelihood and severity by exploiting real-time traffic and weather data collected from urban arterials in Athens, Greece. *Method:* Random Forests (RF) are firstly applied for preliminary analysis purposes. More specifically, it is aimed to rank candidate variables according to their relevant importance and provide a first insight on the potential significant variables. Then, Bayesian logistic regression as well finite mixture and mixed effects logit models are applied to further explore factors associated with accident likelihood and severity respectively. *Results:* Regarding accident likelihood, the Bayesian logistic regression showed that variations in traffic significantly influence accident occurrence. On the other hand, accident severity analysis revealed a generally mixed influence of traffic variations on accident severity, although international literature states that traffic variations increase severity. Lastly, weather parameters did not find to have a direct influence on accident likelihood or severity. *Conclusions:* The study added to the current knowledge by incorporating real-time traffic and weather data from urban arterials to investigate accident occurrence and accident severity mechanisms. *Practical application:* The identification of risk factors can lead to the development of effective traffic management strategies to reduce accident occurrence and severity of injuries in urban arterials.

1. Introduction

The effective treatment of road accidents and thus the enhancement of road safety is a major concern to societies due to the losses in human lives and the economic and social costs. According to World Health Organization (WHO) (2013), the total number of road fatalities worldwide remains at 1.24 million per year. In 2013, some 25,900 people were killed in the European Union because of road accidents, around 313,000 were seriously injured and many more suffered slight injuries (ETSC, 2013). In 2013, 9919 people were killed in traffic accidents on urban roads in the EU, corresponding to 38% of all traffic accident fatalities in 2013 (ERSO, 2015).

Transportation researchers and practitioners have devoted great efforts in order to improve road safety by identifying causes of road accidents (Thomas, Morris, Talbot, & Fagerlind, 2013; Vanlaar & Yannis, 2006). Recently, the progress in technology has enabled the easy recording and collection of real-time traffic and weather data in freeways. Such data were utilized when analysing accident likelihood and severity in freeways. For example, a large number of studies have explored freeway models in order to model accident likelihood (Oh, Oh, Ritchie, & Chang, 2001; Lee, Hellinga, & Saccomanno, 2003; Zheng, Ahn, & Monsere, 2010; Abdel-Aty, Hassan, Ahmed, & Al-Ghamdi, 2012; Xu, Tarko, et al., 2013) and accident severity (Christoforou, Cohen, & Karlaftis, 2010; Yu & Abdel-Aty, 2014a, 2014b).

From the review, it can be seen that there is very little research on safety of urban arterials by using real-time data (Theofilatos & Yannis, in press; Yannis, Theofilatos, Ziakopoulos, & Chaziris, 2014). Although freeway safety has been extensively explored, the transferability of results is questionable. One reason is the completely different environment from an engineering perspective. For instance, freeways have no intersections, no traffic signals, and traffic per direction is fully separated. Moreover, speed limits in freeways are usually 120 km/h. On the other hand, speed limits on urban corridors can vary from 40 km/h to 80 or 90 km/h. The number of lanes are also different in urban arterials than freeways, as urban arterials usually have either one or two lanes per direction. One other important reason is the kind of interactions that take place in urban corridors due to lack of traffic segregation and due to the possibly high percentage of motorcycles and mopeds in traffic. One should also consider the morning and afternoon peak in

E-mail address: atheofil@central.ntua.gr.

http://dx.doi.org/10.1016/j.jsr.2017.02.003

Please cite this article as: Theofilatos, A., Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials, *Journal of Safety Research* (2017), http://dx.doi.org/10.1016/j.jsr.2017.02.003

urban corridors. Lastly, the vast majority of existing studies have a focus on the United States and China. Therefore, existing literature for European road networks is scarce.

In this study, there is focus on accident likelihood and severity on two similar major urban arterials in Athens, Greece. In addition to the traditional accident data obtained from the Greek accident database (SANTRA), real-time traffic and weather parameters are also considered in this study. Accident likelihood refers to the probability that an accident occurs. For that reason, non-accident cases were considered as well (please see data preparation section for the followed research approach). Accidents were classified as severe (accident including at least one seriously injured or killed person) or slight (accident including only slightly injured persons). Due to data limitations (lack of too microscopic measurements[1] of traffic), a more mesoscopic analysis approach was followed, in which the data aggregation was not as microscopic as in previous related studies in the field.

In order to analyse accident likelihood and severity a series of promising statistical methods were applied. Firstly, data mining techniques such as the Random Forests (RF) are applied to rank variable importance and consequently provide a first insight on the significant variables. Then, having acquired information from the Random Forest models, a series of logit models are applied. More specifically, Bayesian logistic regression analysis is carried out to investigate accident likelihood. For comparison purposes, finite mixture models and mixed effects logistic models are applied for modelling accident severity. The results of this study aim to provide an insight on accident likelihood and severity mechanisms and also add to the current knowledge, by including real-time traffic and weather data for urban arterials.

The remainder of the paper is organized as follows. A review of relevant literature is demonstrated, followed by the proposed methodology applied in to model accident likelihood and severity. Then the data description and preparation are provided in Section 3. Next, the application of the models is explained and the results are presented and discussed. The final two sections provide the conclusions as well as the practical applications of the study.

2. Background

Road accidents are resulted from a complex interaction of three fundamental causes: driver, vehicle, and environmental factors. Aside from these factors, other factors may also influence accident occurrence and severity such as socioeconomic factors, legislation, and of course randomness. Consequently, understanding the various factors that cause road accidents and their combined influence is very crucial. Although there has been a very considerable research effort so far, there is still much to be investigated, especially in order to acquire a better knowledge of detailed pre-accident conditions in order to have a better proactive safety management in major roads of the transport network. The advances in the field of Intelligent Transport Systems (ITS) and Meteorology enabled the constant and detailed monitoring of real-time traffic and weather conditions and have contributed to the safety assessment of major roads.

The most frequently adopted approach to explore accident likelihood with real-time traffic and weather data is to include data for accident cases but also for random non-accident cases and apply logistic models (Abdel-Aty & Pande, 2005; Abdel-Aty, Pande, Lee, Gayah, & Dos Santos, 2007; Ahmed & Abdel-Aty, 2012). As accident likelihood analysis is a typical binary classification problem, the most widely applied methods for accident likelihood are the Bayesian logistic regression and the conditional logistic regression. Other models such as neural networks and Support Vector Machines have also been applied (Pande, Q5 Das, Abdel-Aty, & Hassan, 2011; Yu & Abdel-Aty, 2013a, 2013b, 2013c).

Despite the high quality of most studies, contradictory findings are often reported. Kockelman and Ma (2007), found that there was no correlation between 30-sec speed changes and accident likelihood. On the other hand, Ahmed and Abdel-Aty (2012), found that increases in speed variation at the segment of the accident increases the likelihood of accident occurrence. Xu, Tarko, et al. (2013) suggested that accident risk is correlated with high traffic density upstream, increased speed variance, and high differences in volumes and occupancies between upstream and downstream volume detectors.

On the other hand, adverse weather is usually associated with increased risk (Xu, Wang and Liu, 2013, Xu, Tarko, et al., 2013). For Q6 example, Usman, Fu, and Miranda-Moreno (2012) stated that low visibility, increased wind speed and low temperatures are associated with accident risk. Xu, Wang et al. (2013) found that rainfall intensity Q7 is a common risk factor. Ahmed, Abdel-Aty, and Yu (2012), stated that low visibility and high precipitation increase the likelihood of accidents in winter. On the other hand, there are some studies suggest that air temperature and precipitation do not play a significant role in accident occurrence (Usman, Fu, & Miranda-Moreno, 2010).

One important note is that the accident mechanism might not be the same across different time periods. For instance, Yu and Abdel-Aty (2013a) found that weekday crashes are more likely to occur in congested sections, while the weekend accidents are more likely to occur under free flow conditions. Ahmed et al. (2012) investigated the impact of geometrical, traffic, and weather characteristics on accident occurrence on freeways on a mountainous freeway. In winter, it is suggested that low visibility, high precipitation, and speed variation increase the likelihood of accidents, but in dry season, low average speed and low visibility are positively correlated with accident risk. It is also observed that the majority of relevant literature concerns freeways. Only a few studies investigate accident likelihood in urban expressways (Hossain & Muromachi, 2012, 2013). For example, Hossain and Muromachi (2013), aimed to identify accident predictors on urban expressways. One essential contribution is that accident risk in freeway segments and ramp vicinities were analysed separately. The findings suggest that accident mechanism is not the same for basic freeway segments and ramps.

The analysis of accident severity is also of great interest. The key part of such analysis is to understand the way which various contributing factors influence accident severity. Such factors may include driver and passenger attributes, geometric design characteristics, traffic conditions, weather, vehicle type, etc. (Al-Ghamdi, 2002; Chang & Wang, 2006; Milton, Shankar, & Mannering, 2008; Quddus, Noland, & Chin, 2002; Savolainen & Mannering, 2007; Sze & Wong, 2007; Yamamoto & Shankar, 2004; Yau, 2004). Injury severity data are usually expressed by two or more discrete categories, according to the outcome of the accident. Overall, the studies utilizing real-time traffic and weather data to examine accident severity are relatively few (Jung, Qin, & Noyce, 2010; Yu & Abdel-Aty, 2014a, 2014b), while some studies investigate both accident likelihood and severity (Xu, Tarko, et al., 2013).

In Xu, Tarko, et al. (2013) it is suggested that congested traffic leads to less severe accidents because of the lower driving speeds. Another common traffic risk factor is the speed deviation. Yu and Abdel-Aty (2014a, 2014b) found that the standard deviation of speed leads to more severe accidents. Christoforou et al. (2010) investigated injury severity on the A4–A86 junction in the Paris region by applying a random parameters ordered probit model. The authors state that traffic volume is associated with less severe injuries and in addition has fixed effects across observations as no heterogeneity exists (no difference among observations). Another study indicates that traffic conditions had only a marginal influence on accident severity (Golob, Recker, & Pavlis, 2008). However, Yu and Abdel-Aty (2014a), applied Hierarchical Bayesian probit found that large speed variations and low visibility increase accident severity.

Regarding the weather effects on accident severity, the influence of rainfall generally leads to contradictory findings as literature sometimes

[1] Some studies utilize 90 s, 60 s or event 30 s traffic measurements.

Please cite this article as: Theofilatos, A., Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials, *Journal of Safety Research* (2017), http://dx.doi.org/10.1016/j.jsr.2017.02.003