Accident Analysis and Prevention 108 (2017) 172–180

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Full length article

Predicting motorcycle crash injury severity using weather data and alternative Bayesian multivariate crash frequency models

Wen Cheng[a,*], Gurdiljot Singh Gill[a], Taha Sakrani[a], Mohan Dasu[b,1], Jiao Zhou[a]

[a] Department of Civil Engineering, California State Polytechnic University, Pomona, 3801 W. Temple Ave., Pomona, CA, 91768, United States
[b] California Department of Public Health, Sacramento, CA, 95899-7377, United States

ARTICLE INFO

Keywords:
Motorcycle crash
Random parameters
Weather
Rainfall
Temporal
Multivariate severity

ABSTRACT

Motorcycle crashes constitute a very high proportion of the overall motor vehicle fatalities in the United States, and many studies have examined the influential factors under various conditions. However, research on the impact of weather conditions on the motorcycle crash severity is not well documented. In this study, we examined the impact of weather conditions on motorcycle crash injuries at four different severity levels using San Francisco motorcycle crash injury data. Five models were developed using Full Bayesian formulation accounting for different correlations commonly seen in crash data and then compared for fitness and performance. Results indicate that the models with serial and severity variations of parameters had superior fit, and the capability of accurate crash prediction. The inferences from the parameter estimates from the five models were: an increase in the air temperature reduced the possibility of a fatal crash but had a reverse impact on crashes of other severity levels; humidity in air was not observed to have a predictable or strong impact on crashes; the occurrence of rainfall decreased the possibility of crashes for all severity levels. Transportation agencies might benefit from the research results to improve road safety by providing motorcyclists with information regarding the risk of certain crash severity levels for special weather conditions.

1. Introduction

Motorcycle crashes comprise a very high proportion of the overall motor vehicle fatalities in the United States. In terms of the report of the National Highway Traffic Safety Administration (NHTSA, 2016a), in 2015, there were 4976 fatalities on motorcycles in the United States, which represent 14% of the total traffic fatalities. Besides, approximately 88,000 non-fatal injuries resulted from motorcycle traffic collisions. In California, the motorcyclist fatalities increased by 12% from 463 in 2013–519 in 2014 (Office of Traffic Safety, 2014a). Previous studies suggest that motorcyclists represent a group of road users with high crash risks (Clarke et al., 2004; Elliott et al., 2009; Breitenbach et al., 2011) because of their lack of protection in the event of a crash (Rifaat et al., 2012), as well as due to low motorcycle conspicuity or the inability of the vehicle drivers to see the motorcyclists (Hurt et al., 1981; Wulf et al., 1989; Cercarelli et al., 1992). Compared to other transportation modes, the number of serious motorcycle crashes is relatively high, corresponding to the number of motorcycles on the road and driving distance (Haworth et al., 2009; Jama et al., 2011; Bjørnskau et al., 2012). Fatality risk among motorcyclists is about 30 times higher and serious injuries are about 8 times higher compared to other groups of drivers (Huang and Preston, 2004). Specifically, in 2014, the number of motorcyclist fatalities per 100 million VMT (Vehicle Miles Traveled) was more than 23 times that of passenger car crashes (NHTSA, 2016b). As a result, a comprehensive understanding of crash-involvement of motorcyclists is important and much needed.

To this end, a large number of studies have been dedicated to investigating the dependency of motorcycle crashes on various factors including psychosocial factors (Tunnicliff et al., 2012; Tholén et al., 2013), alcohol and drugs (Christophersen and Gjerde, 2015), roadway geometry (Pai and Saleh, 2008), helmet use (Erhardt et al., 2016; Rice et al., 2016), driver behavior (Manan and Várhelyi, 2015), driver age (Yeh and Chang, 2009), and so on. Among them, a subset of studies explored the impact of weather conditions on motorcycle crashes. Due to increased safety risk during adverse weather conditions (Andrey et al., 2001), many studies were focused on the assessment of various weather elements on the vehicle crashes such as precipitation (Andrey and Yagar, 1993; Yu et al., 2013), wind speed (Jung et al., 2011), temperature (Brijs et al., 2008), and winter storms (Knapp, 2000). Comparatively speaking, there are a very limited number of studies which have linked weather effects with

* Corresponding author.
E-mail addresses: fwen26@gmail.com, wcheng@cpp.edu (W. Cheng), gurdiljotg@cpp.edu (G.S. Gill), tsakrani@cpp.edu (T. Sakrani), rkmdasu@gmail.com (M. Dasu), jiaozhou@cpp.edu (J. Zhou).

[1] Affiliation for this author is given for information purposes only. CDPH does not validate or authenticate the views/data presented in the paper and is the sole and only view of the author.

http://dx.doi.org/10.1016/j.aap.2017.08.032
Received 4 May 2017; Received in revised form 17 July 2017; Accepted 29 August 2017
0001-4575/

motorcycle crashes. A study by Pai and Saleh (2008) used ordered logit models for different motorcyclist injury severities of crashes at T-junctions while incorporating variables such as human, road and vehicle factors, and weather conditions (indicated as 'Other', 'Fine', and 'Bad' weather). This study shows that 'fine' weather was observed to be statistically significant and positively related to severe injury crashes. In a subsequent study, Pai (2009) investigated the influence of weather (among other factors) on the failure of motorists to yield to motorcycles. This study used the binary logit model and observed that fine weather conditions positively influenced fatal and severe injury angled perpendicular crashes where a turning car was in a collision with a travelling-straight motorcycle. Another study (Shaheed et al., 2013) employed a mixed logit model to explore the influential factors for severity crashes of motorcycle and another vehicle. The conditions for weather were: Clear, partly cloudy, cloudy, and other. The study observed a non-uniform impact of seasons on minor injury and indicated the unobserved heterogeneity of warm/sunny weather which influenced this impact. Although these studies included weather condition as one among an array of other variables, a comprehensive research study is still needed to explore the influence of weather elements on motorcycle crashes of different injury severities, especially given the increased vulnerability of motorcyclists under adverse weather conditions.

In addition to the exploration of the relationship between various factors and injury severity, another characteristic of the above-mentioned motorcycle crash-related studies is that the proposed models were restricted to only one type of crash at a time. In fact, some researchers (Bijleveld, 2005; Park and Lord, 2007) have revealed the interdependence of different crash outcomes and time periods due to sharing of unaccounted factors and recommended incorporating correlation to obtain more precise parameters for increasing efficiency of a model (Congdon, 2001). Following the recommendation, some studies considered correlations of different crash severity levels through multivariate crash count models (Tunaru, 2002; Miaou and Song, 2005; Ma and Kockelman, 2006; Ma et al., 2008), while others addressed the serial correlations using different treatment of temporal aspects in the crash models (Hay and Pettitt, 2001; Wang et al., 2013; Aguero-Valverde, 2013; Jiang et al., 2014). All these studies noted that more precise estimates were obtained for model parameters with the inclusion of multivariate and serial correlations compared with the competing models without taking into consideration such correlations.

Building upon the previous research studies, the authors aim to conduct a comprehensive research study centered on the investigation of the weather impact on motorcycle crash severities, while addressing multivariate and temporal correlations in the crash data. Specifically, the present study employed five multivariate Poisson lognormal models with different temporal treatments of motorcycle crash data ranging from somewhat simple to more sophisticated ones. It is anticipated that the use of multiple models would not only illustrate the strengths and weaknesses associated with different ways of dealing with the serial correlations, but also yield more reliable findings regarding the influence of weather conditions on motorcycle crash severity levels. Additionally, alternative goodness-of-fit and evaluation criteria were used to assess the model performances from different perspectives.

2. Data description

The focus of this study was the analysis of motorcycle crashes of different severity levels in the city of San Francisco during the years 2008–2013. This city was selected as the area of focus for this study as it has been ranked by the Office of Traffic Safety (OTS, 2014b) as the most unsafe for motorcycles among the cities of California. The crash dataset was obtained from the Transportation Injury Mapping System (TIMS) and comprised of four types of severity based crash outcomes:

- Fatal (FT)
- Serious injury (SI)
- Other visible injury (OI)
- Complaint of pain (CP)

These severity outcomes correspond to the KABCO injury classification scale for the State of California (FHWA). The total number of reported motorcycle crashes of above four severity levels for the time period was 2141, with 41 fatalities, 244 serious injuries, 884 other visible injuries, and 974 cases of complaint of pain in the City. Due to the issue of underreporting related to the property damage only (PDO) crashes (Ye and Lord 2011), they were not incorporated in the crash dataset for model development to avoid potentially biased inferences. The use of four levels of crash severity in the study is insightful since it allows readers to determine whether levels can be aggregated or not. However, the authors were unclear about the criteria used to differentiate severity levels (e.g., fatal vs. serious injury) when the data were initially collected on site. The readers should choose the severity levels or combination thereof based on their best judgement for their future study. To investigate the influence of daily weather changes on the motorcycle crashes, an array of weather-related explanatory variables for the City were obtained from Weather Underground, which updates every hour. Weather Underground observation sources include Personal Weather Stations, National Weather Service, airport data, and weather balloon data. The dataset contained the mean daily Temperature; the total amount of daily Precipitation; the mean Humidity; Cloud Cover; Dew Point; Pressure, among others. In addition, a main exposure-related factor of city safety performance, that is, Daily Vehicle Miles Travel (DVMT), was collected from Highway Performance Monitoring System (HPMS) for the same time periods. It is important to note that the DVMT collected was for all vehicles, which was used to account for the shortage of motorcycle traffic. The similar practice can be found in some of the previous studies (Haque et al., 2010; Flask et al., 2014; Gabauer and Li, 2015) which employed the average daily traffic (ADT) as the exposure predictor. Ideally, motorcycle specific vehicle miles traveled or ADT shall be used if available for less biased estimates of model parameters. Furthermore, a main demographic factor, or, yearly population, was gained from the California Department of Finance. The dependent variable for the models were four severity outcomes of motorcycle crashes while the explanatory variables comprised of weather covariates and day of the week dummy variables.

A statistical summary of the data is presented in Table 1. As evident from the table, the stats for crash data are not aggregated but rather represent the unit observation (one day) considered in this study. It is worth mentioning that, before incorporating the variables for model development, potential insignificant multi-collinear covariates were taken out in steps to ensure that no significant variable was excluded. The correlation tests were conducted using the Harrell Miscellaneous package in R software (Team RC, 2000). The variables observed to be correlated at a significance level of 0.05 were eliminated including yearly population, mean dew point, mean sea level pressure, mean visibility, mean wind speed, the presence of fog and thunderstorm.

3. Methodology

3.1. Model development

This study analyzed four levels of severities for motorcycle crashes occurring in the city of San Francisco. For the development of crash prediction models, Full Bayesian (FB) hierarchical approach was employed to account for the complicated model specifications such as multivariate and temporal correlations, which address the heterogeneities among observed crash severities. Several recent studies have revealed the capability of hierarchical modeling technique to better fit the crash count data by incorporating such heterogeneities (Huang et al., 2008; Aguero-Valverde and Jovanis, 2009; Cheng et al., 2017). FB approach, which is also based on Bayes' theorem, has an advantage over the traditional Empirical Bayesian (EB) method as it can produce a smoother integration of prior