Accident Analysis and Prevention 78 (2015) 146–154

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level

Jaeyoung Lee *, Mohamed Abdel-Aty, Ximiao Jiang

Department of Civil, Environmental & Construction Engineering, University of Central Florida, Orlando, FL 32816-2450, United States

ARTICLE INFO

Article history:
Received 19 June 2013
Received in revised form 13 August 2014
Accepted 3 March 2015
Available online 16 March 2015

Keywords:
Multivariate modeling
Macroscopic analysis
Bayesian modeling
Spatial modeling
Traffic analysis zones
Transportation safety planning

ABSTRACT

Macroscopic traffic crash analyses have been conducted to incorporate traffic safety into long-term transportation planning. This study aims at developing a multivariate Poisson lognormal conditional autoregressive model at the macroscopic level for crashes by different transportation modes such as motor vehicle, bicycle, and pedestrian crashes. Many previous studies have shown the presence of common unobserved factors across different crash types. Thus, it was expected that adopting multivariate model structure would show a better modeling performance since it can capture shared unobserved features across various types. The multivariate model and univariate model were estimated based on traffic analysis zones (TAZs) and compared. It was found that the multivariate model significantly outperforms the univariate model. It is expected that the findings from this study can contribute to more reliable traffic crash modeling, especially when focusing on different modes. Also, variables that are found significant for each mode can be used to guide traffic safety policy decision makers to allocate resources more efficiently for the zones with higher risk of a particular transportation mode.

1. Introduction

Traffic safety is considered one of the most important issues in the transportation field. Although recent data shows both fatalities and fatality rates from traffic crashes in the United States have been steadily declining from 2005 to 2011 but they began to rise again in 2012. Overall 33,561 lives were lost from traffic crashes in 2012 (NHTSA, 2013). Also traffic crashes are one of the leading causes of death in the United States, according to the CDC statistics (Hoyert and Xu, 2012). Accordingly, the Transportation Equity Act for the 21st century (TEA-21) and the Safe, Accountable, Flexible, Efficient, Transportation Equity Act: A Legacy for Users (SAFETEA-LU) proposed the requirement to incorporate safety into the transportation planning process (FHWA, 2005). Thus, traffic safety engineers need to put more efforts to identify crash hotspots and provide appropriate counter measures to reduce the number of crashes accompanied by the long-term transportation plans.

In order to identify crash-prone locations or discover specific contributing factors of the locations for traffic crash occurrences, crash locations are aggregated into spatial units such as segments, intersections, zones, and so forth. Generally, the aggregated safety analysis can be divided into two categories. The first category is micro-level safety analysis which focuses on specific roadway entities such as roadway segments, intersections, corridors, et cetera. The micro-level safety analysis attempts to find out factors affecting traffic crash occurrence from geometric designs and/or traffic characteristics of roadway entities, and suggests specific engineering solutions to reduce traffic crashes. Abdel-Aty and Radwan (2000), Milton and Mannering (1998), Poch and Mannering (1996), and Shankar et al. (1995) examined the traffic safety at the microscopic level. On the other hand, the second category is macro-level, or zonal-level, safety analysis which concentrates on zonal-level traffic crash involvements with socio-demographic contributing factors of the zones. The macro-level provides a broad spectrum perspective, and it suggests long-term policy based countermeasures including enactments of traffic rules, police enforcements, education/campaign, and area-wide engineering treatments. Readers are referred to Abdel-Aty et al. (2013), Lee et al. (2014a,b); Levine et al. (1995), and Quddus (2008) for the better understanding of the macroscopic safety analysis. In this study, we analyzed crashes by different transportation modes at the macroscopic level because we aimed at providing guidance to

* Corresponding author. Tel.: +1 407 823 2733.
E-mail address: jaeyoung@knights.ucf.edu (J. Lee).

http://dx.doi.org/10.1016/j.aap.2015.03.003
0001-4575/

policy decision makers to effectively reduce traffic crashes by various modes along with the long-term transportation planning process.

In the methodological point of view, multivariate modeling per se is not a novel method for analyzing multiple crash types. However, not many studies have applied multivariate models for the macroscopic safety analysis. In addition, no macro-level studies have utilized multivariate models for crashes by different modes so far. If there are strong correlations in errors across mode types it implies the existence of common unobserved factors across crash modes. In this case, the multivariate can handle the shared unobserved factors because it includes common error components among various crash types in the model specification. Therefore, the objective of this study is to develop a multivariate model for crashes by transportation modes such as motor vehicle, bicycle, and pedestrian crashes at the macroscopic level. Moreover, we used not only roadway and traffic variables but also demographic and socioeconomic variables which are generally used for long term transportation plans. Thus traffic analysis zone (TAZ) based multivariate models estimated in this study can be more usefully utilized by transportation planners.

2. Literature review

In recent decades, a large body of literature have investigated traffic safety propensity on the macroscopic level, such as block group (Levine et al., 1995; Abdel-Aty et al., 2013), TAZ (Siddiqui, 2009; Washington et al., 2010; Naderan and Shahi, 2010; Abdel-Aty et al., 2011, 2013; Dong et al., 2014), traffic safety analysis zones (Lee et al., 2014a,b), census tract (Loukaitou-Sideris et al., 2007; Wier et al., 2009; Cottrill and Thakuriah, 2010; Ukkusuri et al., 2011; Abdel-Aty et al., 2013), county (Aguero-Valverde and Jovanis, 2006; Amoroset al., 2003; Huang et al., 2010; Noland and Oh, 2004), state (Noland, 2003), ZIP code (Lee et al., 2013, 2014a,b, 2015), and others (Noland and Quddus, 2004; MacNab, 2004; Kim et al., 2006). Among these spatial units, TAZ has been widely adopted for traffic safety analysis. A TAZ is a statistical entity delineated by state Department of Transportation (DOTs) and/or local Metropolitan Planning Organizations (MPOs) officials for tabulating traffic-related census data such as, journey-to-work and place-of-work statistics (U.S. Census Bureau, 2011). Therefore, from a transportation planning perspective, TAZs seem to be preferred spatial entities as compared to other spatial units.

Previous macroscopic safety studies have investigated the occurrence of crashes by various classifications from different perspectives, such as crashes by pedestrian and/or bicycle crashes (Noland and Quddus, 2004; Kim et al., 2006; Abdel-Aty et al., 2011; Siddiqui et al., 2011; Abdel-Aty et al., 2013; Lee et al., 2013), crashes by injury severity levels (Hadayeghi et al., 2003; Noland, 2003; Noland and Oh, 2004; Hadayeghi et al., 2006; Aguero-Valverdeand Jovanis, 2006 Hadayeghi et al., 2010; Naderan and Shashi, 2010 Huang et al., 2010; Abdel-Aty et al., 2013), or crashes during specific time periods (Hadayeghi et al., 2003, 2006; Abdel-Aty et al., 2011). Among these viewpoints, crashes by specific transportation modes such as motor vehicles, bicycles and pedestrians have been investigated by several researchers.

From a methodology perspective, a wide spectrum of modeling approaches has been incorporated in the macro-level safety research so far. The conventional methods include NB models (Hadayeghi et al., 2003, 2006; Siddiqui, 2009; Noland and Quddus, 2004; Karlaftis and Tarko, 1998; Amoros and Laumon, 2003; Noland and Oh, 2004; Aguero-Valverde and Jovanis, 2006), ordinary least square regression model (Wier et al., 2009), log-linear models (Washington, 2006), Bayesian hierarchical models (Quddus, 2008) and Bayesian models accounting for spatial autocorrelation (Huang et al., 2010; Wang et al., 2012; Lee et al., 2015).

Nevertheless, only a few researchers considered multivariate modeling for the macroscopic analysis even though they look into more than one target crash types. As has been proven by many research studies, there is strong correlation among crash frequencies of different types (i.e., severity levels) within each site (Tunaru, 2002; Ma and Kockelman, 2006; Aguero-Valverde and Jovanis, 2009; Park and Lord, 2007; Ma et al., 2008; El-Basyouny and Sayed, 2009). This correlation is caused by some site specific unobserved factors that may affect traffic safety. Similarly, this type of correlation may also exist in macro-level crash frequencies. Ignoring these correlations may lead to biased parameter estimates and thus the corresponding crash frequency prediction (Ye et al., 2009; Ma et al., 2008). In order to address this problem multivariate models have been proposed and widely used in micro-level crash analysis (Tunaru, 2002; Ma and Kockelman, 2006; Park and Lord, 2007; Ma et al., 2008; El-Basyouny and Sayed, 2009; Ye et al., 2009). For example, Ma and Kockelman (2006) applied a multivariate Poisson regression model using crash data from roadway segments in Washington State. It was found that there was positive correlation in unobserved factors affecting crash frequencies across severity levels. Also goodness-of-fit measures showed that the multivariate Poisson model is superior to the suite of independent models. Ye et al. (2009) looked into crashes by collision types such as head-on, sideswipe, rear-end, and angle crashes at intersections. The authors developed univariate and multivariate Poisson regression models and compared the two model structures. No significant differences were found in magnitude of coefficients between the two model systems. Nevertheless, with respect to goodness-of-fit measures (adjusted likelihood ratio index), the multivariate model showed better fit than the univariate model. In addition the authors found that the results revealed the presence of common unobserved factors across collision types. Although, many studies have been done in the micro-level crash analysis, few studies have addressed the potential correlations between each type of crashes in the macroscopic safety analysis. Guevara et al. (2004) utilized simultaneous estimation of the models for injury and fatal crashes using TAZ crash data. Although the authors were not successful to show simultaneous models outperformed the independent models in terms of goodness-of fits, they found there is a significant correlation across disturbance terms of injury and fatal models.

At the same time, few researchers have reported that there is a spatial correlation for macro-level crashes. Spatial autocorrelation is a technical term for the fact that spatial data from near locations have higher probabilities to be similar than data from distant location (O'Sullivan and Unwin, 2002). The existence of spatial autocorrelations in the data may invalidate the assumption of the random distribution (LeSage and Pace, 2004). Several researchers have accounted for spatial effects in the crash model (Levine et al., 1995; Hadayeghi, 2010 Quddus, 2008; Huang et al., 2010; Siddiqui and Abdel-Aty, 2012). LaScala et al. (2000) discovered that there is a significant spatial relationship between pedestrian injury crashes and specific environmental and demographic characteristics of San Francisco. Huang et al. (2010) showed that the variation accounted for by spatial clustering are essential for crash risk models. The authors discovered that spatial autocorrelations were significant in traffic crashes across adjacent counties in Florida. Siddiqui and Abdel-Aty (2012) developed Bayesian Poisson log-normal models with spatial error terms for bicycle and pedestrian crashes in Florida. The authors revealed that models accounting for spatial autocorrelations showed better performance compared to models without considering spatial effects.

Previous papers have significantly contributed to addressing the correlations among crash types, as well as the potential spatial