Accident Analysis and Prevention 132 (2019) 105286

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Modeling of time-dependent safety performance using anonymized and aggregated smartphone-based dangerous driving event data

Di Yang[a], Kun Xie[b,*], Kaan Ozbay[c], Hong Yang[d], Noah Budnick[e]

[a] Department of Civil and Urban Engineering, Connected Cities for Smart Mobility towards Accessible and Resilient Transportation (C2SMART) Center, New York University, 15 MetroTech Center, 6th Floor, Brooklyn, NY 11201, USA
[b] Department of Civil and Environmental Engineering, Old Dominion University, 4635 Hampton Boulevard, Norfolk, VA 23529, USA
[c] Department of Civil and Urban Engineering, Connected Cities for Smart Mobility towards Accessible and Resilient Transportation (C2SMART) Center, and Center for Urban Science and Progress (CUSP), New York University, 15 MetroTech Center, 6th Floor, Brooklyn, NY 11201, USA
[d] Department of Computational Modeling and Simulation Engineering, Old Dominion University (ODU), 4700 Elkhorn Ave, Norfolk, VA 23529, USA
[e] Data Practice & Policy Director (formerly with Zendrive), Zendrive Inc, 929 Market St, San Francisco, CA 94103, USA

ARTICLE INFO

Keywords:
Dangerous driving events
Connected vehicles
Safety performance functions
Time-dependent hotspots
Multivariate conditional autoregressive model

ABSTRACT

Safety performance functions (SPFs) are generally used to relate exposure to the expected number of crashes aggregated over a long time (e.g. a year) by holding all other risk factors constant, and to identify hotspots that have excessive crashes regardless of different time periods. However, it is highly likely that the relationships of exposure, risk factors and crash occurrence can vary across different times of day. This study aims to establish time-dependent SPFs for urban roads by using large-scale dangerous driving event data captured by smartphones in different times of day. Multivariate conditional autoregressive (MVCAR) models are developed to jointly account for spatial and temporal dependence of crash observations. Results of two-sample Kolmogorov-Smirnov tests affirm the heterogeneity of the safety effects of dangerous driving events in different time periods. Time-dependent hotspots are identified using potential for safety improvement (PSI) metric. The assumption here is that due to the change of traffic conditions and environment across different times of day, safety hotspots for different time periods should be different from each other. According to the results of Wilcoxon signed-rank tests, hotspots identified by times of day are found to be mostly different from each other. The findings of this study provide insights into temporal effects of risk factors and can support the development of time-dependent safety countermeasures. Besides, this study also shows the potential of leveraging anonymized and aggregated dangerous driving data to assess traffic safety issues.

1. Introduction

Safety performance functions (SPFs) are commonly used to correlate geometric, traffic, and environmental characteristics with crash outcomes. These models are frequently sought to support the detection of hotspots that have excessive crashes over similar sites (Wang et al., 2014).

It should be pointed out that SPFs are primarily designed to relate exposure to expected number of crashes, usually per year, at a location by holding all other risk factors constant (Hauer, 1995). Such aggregated functions cannot clearly account for the potential variations of crash occurrences at different times of day. Obviously, factors such as traffic exposure and some roadway conditions have distinct patterns at different periods of a day (e.g., congestion at peak hours vs free flow at night). Arguably, this naturally leads to different levels of crash risk at different periods. For example, one may see more speeding events when traffic is less congested.

A number of studies have illustrated the non-uniform distributions of crashes across different times of day (e.g., Qin et al. (2006) and Pahukula et al. (2015)). From the implementation perspective, safety countermeasures should be designed accordingly. However, if conventional SPFs were used, it would be difficult to capture the appropriate association between crashes and time-sensitive risk factors. Consequently, there will be fewer chances to unveil the facts that a location having excessive crashes during certain periods may be relatively safe in the rest of the day. In other words, hotspots can be time-dependent. With a similar concept, Folkard (1997) proposed the term “black times” to indicate high-risk time periods against “black spots” that represent

* Corresponding author.
E-mail addresses: dy855@nyu.edu (D. Yang), kxie@odu.edu (K. Xie), kaan.ozbay@nyu.edu (K. Ozbay), hyang@odu.edu (H. Yang), noah@zendrive.com (N. Budnick).

https://doi.org/10.1016/j.aap.2019.105286
Received 13 February 2019; Received in revised form 31 July 2019; Accepted 26 August 2019
Available online 02 September 2019
0001-4575/

fixed high-risk locations.

To model time-dependent safety performance, it is essential to collect risk factors during different times of day. Such studies are rarely performed because time-dependent risk factors are difficult to obtain using traditional data collection method. Other than relating short-term traffic flow measures with crash occurrences (e.g. Golob and Recker (2004); Qin et al. (2006), and Pahukula et al. (2015)), it is very difficult to gather the many crash precursors, such as dangerous driving events, in urban areas that have been found to be positively correlated with crash or near-crash occurrence (Guo and Fang, 2013; Paleti et al., 2010).

For collecting the dangerous driving event data, one possible solution is to resort to connected vehicle (CV) technologies. CVs are vehicles that are able to communicate with each other (V2V) and with the infrastructure (V2I) through wireless communication technologies (Feng et al., 2015). The exchange of information between connected vehicles can be used to detect risk or potential collisions (Xie et al., 2018ab) and warnings could then be triggered and issued to drivers accordingly (Doecke et al., 2015). For example, if a vehicle hard brakes in front of the host vehicle, then an emergency electronic brake lights (EEBL) warnings will be issued to the host vehicle (Howe et al., 2016). In this sense, the warnings themselves will be a good indicator for crashes.

However, because of the high implementation cost of the CV devices, large-scale installment of CV devices may not be possible shortly. And due to privacy and security concerns, vehicle and driver IDs and other sensitive information that can link travel trajectories to individual vehicle or driver will be scrubbed in future CV data (e.g. in the Connected Vehicle Pilot Deployment (CVPD) Program – New York City (NYC) DOT Pilot (Galgano et al., 2016)). All of these motivate us to explore other data sources that can be served as a substitute to future CV data. Many crash surrogates were proposed in traffic safety literature. For example, time-to-collision (TTC) (Hayward, 1972), deceleration rate to avoid collision (DRAC) (Cooper and Ferguson, 1976), and post-encroach time (PET) (Allen et al., 1978) are among commonly used surrogates for approximating crashes. But, calculating these surrogates often requires the detailed trajectory information of two consecutive vehicles in a car following scenario, which will be difficult to obtain in the real world. Thus, indicators that can be derived from individual vehicle trajectory will be more practical and are the focus of this study. Anonymized and aggregated dangerous driving event data collected through smartphones by a private company Zendrive (Zendrive, 2018) will be used in this study. Driving event data is collected automatically and passively from smartphone sensors while the Zendrive application is active. Specifically, four types of dangerous driving events data that have been shown to be good indicators for crashes in literature (please refer to the Literature Review section) are collected, which are fast acceleration, hard braking, speeding, and phone use while driving. Zendrive records the location and time of each dangerous driving event, which enables the development of time-dependent safety models. Despite the exclusion of vehicle-to-vehicle communication component, Zendrive has very large coverage due to ubiquitous presence of smartphones. This offers the opportunities to probe the safety issues with unique data in a large area rather than individual sites.

This paper aims to develop the time-dependent SPFs by leveraging the large-scale dangerous driving event data along with geometric, traffic, and environmental characteristics. The unique dataset provided by Zendrive allows us to fully explore the safety effects of dangerous events during different times of day, which has been rarely studied in the literature. Manhattan, New York City is used as our study area. A multivariate conditional autoregressive (MVCAR) model, that can account for both spatial and temporal dependences is developed to model time-dependent safety performance of different areas. With the developed model, temporal effects of risk factors are investigated and time-dependent hotspots are identified. Potential time-dependent safety countermeasures are also discussed along with applications of the developed model.

2. Literature review

2.1. Time-of-day-related safety performance

Existing studies have discussed the temporally variable association of crash occurrence with respect to risk factors on road environment, traffic characteristics, and driver behavior (Ivan et al., 2000; Lenné et al., 1997; Maze et al., 2006; Plainis and Murray, 2002). Environmental attributes are critical risk factors that have proven to affect crash rates (Balagh et al., 2014). Two commonly examined road environmental factors are weather and visibility. Maze et al. (2006) reviewed a large number of literature and concluded that in most studies crash rates increased as roadways became wet or covered with snow and/or ice. Obviously, weather conditions are often different at different times of day (e.g., different temperature at different periods). Visibility conditions also vary by times of day. Plainis and Murray (2002) showed that a large number of crashes at night was primarily caused by visibility issues regarding low luminance conditions. Arguably, due to the changes of these variable weather factors, traffic safety performances are also time-dependent.

Other than environmental factors, previous studies have also investigated the temporal effect of traffic characteristics on safety performances of the road network. For example, hourly traffic volumes were used by Ivan et al. (2000) and Martin (2002) when predicting crash counts on highways. Likewise, several studies (Marquis and Wang, 2015; Xie et al., 2015; Zou et al., 2016) have built time-dependent SPFs for truck involved crashes in New York City during different times of day. Traffic volume for each time period was used for crash modeling. Abdel-Aty and Abdalla (2004) modeled freeway crash outcomes at different times of day and found that standard deviation of traffic volume, mean and standard deviation of speed had significant impact on safety performance.

Moreover, driver behavior at different time periods of a day has been explored as well in a limited number of studies. For example, Lenné et al. (1997) performed a driving simulator study and found that mean and standard deviation of lateral position and speed and reaction time were affected by times of day. Folkard (1997) found out that there was a pronounced circadian rhythm in the probability of a crash. Due to the difference of driver behaviors at different times of a day, crash risks imposed by drivers would vary by time periods too.

2.2. Dangerous driving event data for traffic safety research

The relationship between dangerous driving events and traffic crashes has been studied by traffic safety researches (Ryder et al., 2018). Traditionally, dangerous driving event data was extracted from driving data that was collected using devices installed in vehicles, such as on-board devices (OBDs) or in-vehicle data recorders (IVDRs) (Baltusis, 2004; Dingus et al., 2006a; Toledo et al., 2008). The use of on-board driving data requires additional investment in deploying the OBDs, which may not be economic. In addition, the acquisition of the recorded data will be challenging and the data will be limited due to the enrollment of vehicles in such data collection experiments, which may result in low market penetration rate. Recently, with the rapid advance of mobile sensors and technologies, smartphones have gained increasing popularity in driving data collection. Smartphone sensors including GPS receivers (providing location information), gyroscopes (measuring orientation and angular velocity), and accelerometers (measuring acceleration) (Kanhere, 2011) can be used to generate driving data. Such massive high-resolution driving data offer the valuable opportunities to detect and monitor dangerous driving events if appropriate learning algorithms were used (Gitelman et al., 2018; Pande et al., 2017). Comparing to OBDs and IVDRs, the advantages of smartphone-powered data collection method are: a) cost effective

(Fazeen et al., 2012); b) easily accessible and portable (Fazeen et al., 2012); and c) wide coverage and highly deployable in urban areas (Kanhere, 2011).

There were a few studies that have mined the dangerous driving event data for traffic safety analysis. Among early practices, many studies have leveraged naturalistic driving data to extract crash surrogates from vehicle trajectories. Dingus et al. (2006b) and Wu and Jovanis (2011) used longitudinal and lateral accelerations collected from naturalistic driving studies to extract hard braking and fast acceleration among other crash surrogates involving vehicle interactions. (Lee et al., 2011) collected vehicle movement data from traffic simulator to investigate alcohol-related impairment. Other than collecting data in naturalistic driving environment and simulators, multiple smartphone sensors were used by Johnson and Trivedi (2011) to distinguish multiple types of aggressive driving behavior including hard brakings and fast accelerations. Fazeen et al. (2012) identified fast acceleration and hard braking events by filtering excessive longitudinal acceleration and Dunlop et al. (2016) collected hard braking events using smartphones and identified dangerous road sections with excessive hard braking events. Stipancic et al. (2018) proposed a new algorithm to extract hard braking and fast acceleration from vehicle trajectories collected by smartphones. Besides hard brakings and fast accelerations, speeding and phone use while driving behaviors have been explored in literature as well. Delhaye (2006) studied the relationship between crashes and vehicle speed through examining the regulation of speed limit. Speeding behavior was studied through both surveys (Broughton et al., 2009; De Pelsmacker and Janssens, 2007; Elliott et al., 2003) and analyzing naturalistic driving data (Richard et al., 2017, 2012). Specifically, (Richard et al., 2012) examined different types of speeding behaviors among individual drivers. Phone use while driving was explored mainly using two types of study methods, i.e. surveys (White et al., 2010; White et al., 2004) and experimental design (Klauer et al., 2014; Redelmeier and Tibshirani, 1997). As summarized by (J. D. Lee, 2008), phone use increased drivers' reaction time, which may contribute to 12% to 25% of crashes that are attributed to driver distraction or inattention.

After extracting dangerous driving events, the relationship between them and crashes was explored in a few studies lately. The statistical relationship between crash surrogates and crashes was studied by Guo et al. (2010) and Wu and Jovanis (2011) and positive relationships were found between them under various conditions (such as different weather, road surface conditions and so on). Stipancic et al. (2018) examined the association between hard braking and fast acceleration with crash records using Spearman's correlation coefficient and pairwise Kolmogorov-Smirnov test. Both dangerous events were shown to be positively correlated with crash frequencies. Ryder et al. (2018) demonstrate the relationship between vehicle jerk rate and crash rate on a national level using data collected from 72 professional drivers within 18-week. Pande et al. (2017) and Gitelman et al. (2018) both incorporated dangerous driving events into building SPFs. Pande et al. (2017) used vehicle jerk (derivative of acceleration with respect to time) collected from 33 drivers with each driver providing about 10 days of driving data for crash modeling, while braking and speeding alerts collected from 64 drivers in a one year period were included in the SPFs developed by Gitelman et al. (2018). Statistically significant associations were found between dangerous driving events and crashes in both studies. However, because each driver only provides about 10 days of driving data in Pande et al. (2017), it may not be appropriate for building SPFs. Further, despite these efforts, none of the aforementioned studies used dangerous driving event data to model time-dependent safety performance, which will be the focus of this study.

3. Methodology

To evaluate time dependent safety performance, each day was divided into four time periods: a.m. peak (6 a.m. to 10 a.m.), midday (10 a.m. to 4 p.m.), p.m. peak (4 p.m. to 7 p.m.), and night (7 p.m. to 6 a.m.). This temporal separation is consistent with our previous study that explored the safety impacts of truck off-hour delivery programs in New York City (NYC) (Xie et al., 2015). As for the spatial unit of analysis, census tracts (n = 282) of Manhattan were used as the basic geographical units (shown in Fig. 1). Census tracts can be easily connected to the demographic and economic features, and spatial aggregation at the census tract level has been frequently used as the unit of analysis (Abdel-Aty et al., 2013; Aguero-Valverde, 2013b; Quddus, 2008).

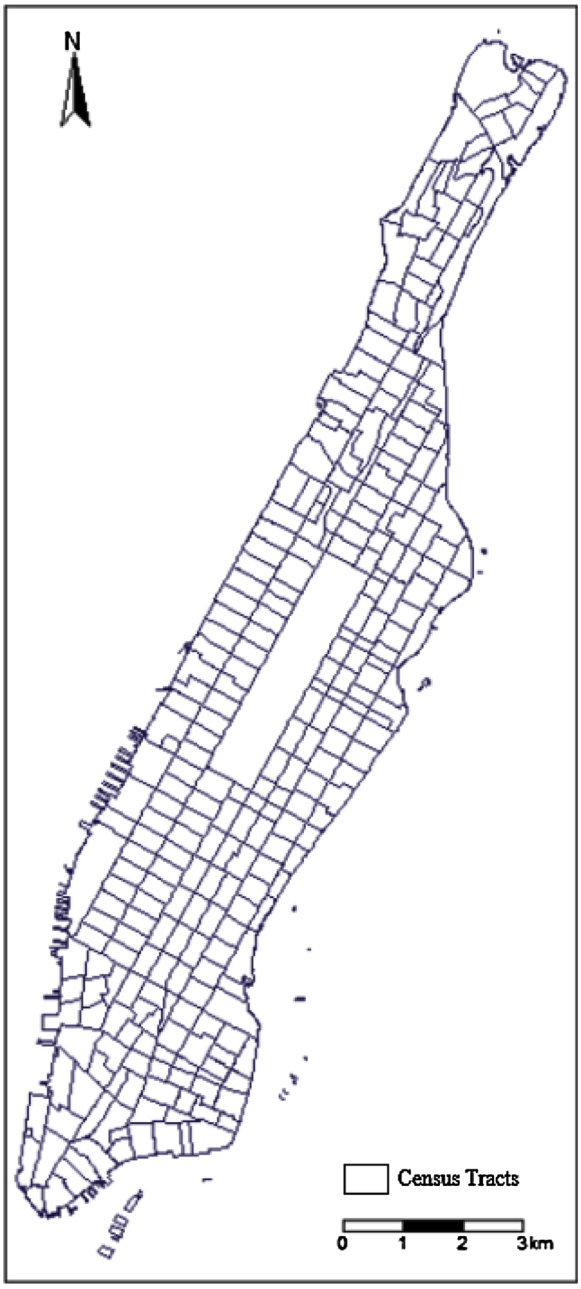

Fig. 1. Census tracts of Manhattan, New York.

3.1. Multivariate conditional autoregressive model

To develop time-dependent SPFs, it is essential to consider the temporal correlation (Zou et al., 2016) among crash frequencies observed in different time periods. Meanwhile, the crash observations of adjacent sites may be spatially correlated (Lord and Mannering, 2010).

Therefore, this study proposes a multivariate conditional autoregressive (MVCAR) model to jointly account for both spatial and temporal dependences. It is built upon the basic Poisson-Gamma (PG) model (Aguero-Valverde, 2013a) and is a multivariate extension of univariate conditional autoregressive (UCAR) model (Aguero-Valverde, 2013b; Quddus, 2008). A K-dimensional multivariate conditional autoregressive (MVCAR) model is specified in Eq. (1).

$$\begin{aligned} y_i^k &\sim Poisson(\lambda_i^k) \\ \ln(\lambda_i^k) &= \beta_0^k + \sum_{p=1}^{P_k} \beta_p^k X_{pi}^k + \varepsilon_i^k + S_{ki} \end{aligned} \tag{1}$$

where y_i^k denotes the crash count in the k^{th} time period at the i^{th} site ($i = 1, 2, ..., n$, n is the total number of sites; $k = 1, 2, ..., K$, K is the total number of time periods). λ_i^k denotes the mean of the Poisson distribution. X_{pi}^k represents explanatory variables, where $p = 1, ..., P_K$ and P_K is the total number of explanatory variables for k^{th} time period. β_0^k and β_p^k are the regression coefficients to be estimated. Error term ε_i^k is included to address the over-dispersion issue, with $\exp(\varepsilon_i^k)$ assumed to be Gamma-distributed with mean 1 and variance σ_k^2. S_{ki} is a multivariate CAR effect term, which can account for the spatial correlation among different time periods and census tracts. The full conditional distribution of $\mathbf{S_i} = (S_{1i}, S_{2i}, ..., S_{Ki})'$ follows a K-dimensional multivariate normal distribution (Thomas et al., 2004):

$$\mathbf{S_i}|\mathbf{S_{-i}} \sim MVN_K\left(\sum_{i' \neq i} \frac{w_{i'i}}{w_{i+}}\mathbf{S}_{i'}, \frac{\mathbf{\Omega}}{w_{i+}}\right) \tag{2}$$

where $\mathbf{S_{-i}}$ is the set of $\mathbf{S}_{i'}$ for any $i' \neq i$. $w_{i'i}$ indicates the spatial autocorrelation between sites i' and i, with $w_{i'i} = 1$ if sites i' and i are adjacent, and $w_{i'i} = 0$ otherwise. w_{i+} is the aggregation of weights for site i, with $w_{i+} = \sum_{j=1}^{n} w_{ij}$. $\mathbf{\Omega}$ is the variance-covariance matrix for specifying correlation structure among different time periods:

$$\mathbf{\Omega} = \begin{pmatrix} \sigma_{S11}^2 & \sigma_{S12}^2 & ... & \sigma_{S1K}^2 \\ \sigma_{S21}^2 & \sigma_{S22}^2 & ... & \sigma_{S2K}^2 \\ ... & ... & ... & ... \\ \sigma_{SK1}^2 & \sigma_{SK2}^2 & ... & \sigma_{SKK}^2 \end{pmatrix} \tag{3}$$

Diagonal elements of $\mathbf{\Omega}$ (i.e., $\sigma_{S11}^2, \sigma_{S22}^2, ..., \sigma_{SKK}^2$) indicates the conditional variance of the spatial effects of individual time period and the off-diagonal elements (i.e., $\sigma_{S12}^2, \sigma_{S13}^2, ..., \sigma_{SKK-1}^2$) represent the conditional with-in site covariance of the spatial effects of different time periods. For detailed discussion of PG model, UCAR model, and MVCAR model specifications, please refer to (Thomas et al., 2004).

3.2. Bayesian method

The proposed MVCAR model is estimated in the full Bayesian framework. Integrating with the Bayesian method provides us advantages to flexibly select crash count distributions, and to accommodate complicated model structures (Lan et al., 2009; Mitra and Washington, 2007; Xie et al., 2013). Bayesian inference generally is performed using Markov Chain Monte Carlo (MCMC) algorithm (Gilks et al., 1998). The WinBUGS statistical software package was used to provide a computing approach for the calibration of Bayesian models using MCMC simulation (Spiegelhalter et al., 2002a).

The priors of the multivariate CAR effect terms $\mathbf{S_i}$ were generated from Eq. (2). All regression coefficients priors were assumed to follow the Gaussian distribution $(0,10^2)$. The variance of the univariate CAR distribution σ_{Sk}^2 were assumed to follow the Inverse-Gamma distribution $(10^{-3}, 10^{-3})$. The logarithm of the variance of the Poisson-Gamma error term $(\ln \alpha_k)$ was assumed to follow the Gaussian distribution (0, 10). The variance-covariance matrix for correlation $\mathbf{\Omega}$ was assumed to follow a Wishart distribution (Thomas et al., 2004).

Potential scale reduction factor (PSRF) (Gelman and Rubin, 1992) was used to check the convergence of multiple Markov chains. PSRF is based on a comparison of within-chain and between-chain variances. As the number of iteration becomes larger, PSRF declines to 1 if the algorithm converges. In this study, convergence was assumed to occur when PSRF is less than 1.1, which is consistent with previous practice in Sinharay (2003). The number of burn-in for each model is tuned accordingly.

For model performance evaluation and comparing competing models, Deviance Information Criterion (DIC) is used in this study (Spiegelhalter et al., 2002a). Please refer to (Spiegelhalter et al., 2002b) for a detailed discussion of DIC. A difference in DIC that is larger than 5 suggests that the model with a smaller DIC should be favored (Spiegelhalter et al., 2003).

3.3. Potential for safety improvement

Given time and budget constraints, only a portion of the census tracts should be identified as hotspots to implement countermeasures (Xie et al., 2017). The naïve methods that simply rely on the raw crash observations such as crash frequencies and crash rate are among the early practice of hotspot identification (Xie et al., 2017). A well-known limitation of the naïve methods is the regression-to-the-mean (RTM) issue. Since crashes are rare and random events, sites flagged as hotspots due to high crash frequency in one period can experience lower crash frequencies subsequently even no countermeasures is implemented (Hauer, 1997). Potential for safety improvement (PSI) is used widely in literature in addressing RTM issue (Elyasi et al., 2016; Hauer et al., 2002; Persaud et al., 1999; Xie et al., 2017). Specifically, PSI can be defined as the expected crash count of one site minus the estimated crash count of "similar" sites. The safety effects of exposure indicators (e.g., VMT) can be accounted for in the crash frequency models and thus PSI can capture the portion of crashes which are caused by unobserved site-specific risk factors.

The formula for calculating PSI based on MVCAR model is given in Eq. (4).

$$PSI_i^k = \exp(\beta_{0i}^k + \sum_{p=1}^{P_k} \beta_{pi}^k X_{pi}^k + \varepsilon_i^k + S_{ki}) - \exp(\beta_{0i}^k + \sum_{p=1}^{P_k} \beta_{pi}^k X_{pi}^k) \tag{4}$$

where PSI_i^k is the potential for safety improvement of the k^{th} time period at site i. $\exp(\beta_{0i}^k + \sum_{p=1}^{P_k} \beta_{pi}^k X_{pi}^k + \varepsilon_i^k + S_{ki})$ represents the expected crash count of the k^{th} time period for site i, which can be estimated using Eq. (1). $\exp(\beta_{0i}^k + \sum_{p=1}^{P_k} \beta_{pi}^k X_{pi}^k)$ represents the average crash count for sites which have similar attributes to site i in the k^{th} time period.

4. Data preparation and exploration

Two major data sources used in this study are crashes and dangerous driving events. Other data sources, including transportation data, land use data, and socio-demographic data, are also collected for analysis.

4.1. Dangerous driving event data

As mentioned in the Introduction, dangerous driving event data used in this study was provided by Zendrive, a private company launched in 2013 and based in San Francisco. The dangerous driving event data collected by their system have been used to explore various traffic safety issues, such as distracted driving (mainly, using phones while driving events) issue near campuses (Sutton, 2017).

Four types of dangerous driving events are used in this study, namely fast acceleration, hard braking, speeding, and phone use while driving. The 6-month dataset collected covered the days from July 1st, 2015 to December 31st, 2015. The total number of events is 10,512. Each type of events was reported based on Zendrive's internal algorithms and criteria that typically consider the excessive amount of deceleration, acceleration, prevailing speed, and actions related to phone

use, respectively. According to Zendrive, a hard braking event is identified if a reduction in the speed is fast enough to thrust the driver and passengers' bodies forward hard enough to cause the seatbelt to lock. A fast acceleration event is identified if the acceleration is fast enough to throw the driver and passengers' bodies back in their seats. A speeding event is identified when a driver is going above the posted speed limit by a set amount for at least 10-seconds. A phone use while driving event is identified when the driver is traveling over a minimum speed and uses their phone in their hand for at least 3-seconds. Fig. 2 provides an overview of the distributions of these events. According to Fig. 2(a), speeding is the most frequent event that is about three times more than the fast acceleration events. The monthly distribution in Fig. 2(b) shows that the proportions of speeding and fast acceleration events are relatively stable in each month, whereas the proportions of the other two types of events change monthly. Fig. 2(c) illustrates the distributions of four types of dangerous driving events during different periods of a day. Most dangerous driving events occur at night but less during both a.m. and p.m. peak periods.

Because dangerous event data was collected by smartphones and our study region is Manhattan, it is necessary to examine the quality of the GPS location because GPS signals might be affected in urban canyons. In this study, dangerous event points are mapped to Manhattan street network with a 30 feet buffer. 89.6% points can be mapped, which indicates a good GPS signal quality. As an exploratory analysis, the spatial distributions of the dangerous driving events broken down by time periods are shown in Fig. 3(a). The visualization represents the kernel density maps created using the ggmap package of R software (Kahle and Wickham, 2013). As shown in Fig. 3(a), the high-density areas of dangerous driving events at different time periods are distinguishable though most of the points are located at downtown and midtown Manhattan. Regardless of time periods, more events are also clustered nearby the entrances of bridges and tunnels. More points are located at the east side of Manhattan at night.

4.2. Crash data

To be consistent with dangerous driving event data, half-year crash data (07/01/2015 – 12/31/2015) were obtained from the NYC Open Data website[1] . A total of 23,758 crashes occurred in Manhattan were identified. The proportions of crashes occurred at the four study time periods (a.m. peak, midday, p.m. peak, and night) are 13.3%, 35.5%, 20.0%, and 31.2%, respectively.

As an exploratory analysis, the kernel density maps for crashes occurred in each time period are shown in Fig. 3(b). Comparing to Fig. 3(a), the general spatial distributions of crashes and dangerous driving events shared some similarity, e.g., most points are clustered around downtown and midtown Manhattan. But specifically, the major cluster of crashes is located around the center of midtown Manhattan. The number of crashes is the highest at night. Quantatively, the Pearson correlation coefficients were calculated between the crashes and dangerous driving events at each period among the studied census tracts. The obtained Pearson correlation coefficients for a.m. peak, midday, p.m. peak and night are 0.42, 0.47, 0.51, and 0.49, all of which suggest a statistically significant correlation between crashes and the dangerous driving events. Corresponding scatterplots between the log of the number of crashes and the log of the number of dangerous driving events are shown in Fig. 4 along with the linear regression lines. As can be seen from the figure, there is a roughly linear relationship between number of crashes and number of dangerous driving events for each time period. This indicates that the relationship between the log of the number of dangerous driving events and the log of crashes for each time period is also roughly linear but with the increase of number of dangerous events and crashes for each plot, the variance increases too which results in several outliers.

4.3. Other data used in this study

Traffic AADT data were obtained from the Short Count Program (SCP) of NYSDOT.[2] In the SCP, approximately 12,000 statewide counts of 2–7 days' duration were taken every year and were used to calculate AADT after undergoing quality control procedures. Vehicle miles traveled (VMT) was computed for each census tract based on the AADT data. The output of the Best Practice Model (BPM) developed by the New York Metropolitan Transportation Council (NYMTC[3]) was used to estimate the ratio of truck flow to total flow for each census tract. The geographic information system (GIS) data of bus and subway stations were obtained from the Metropolitan Transportation Authority (MTA[4]). The number of bus and subway stations was calculated for each census tract using spatial tools of the software ArcGIS. The intersection number and length of different roadways were computed based on the road network of the regional planning model.

The land use data were obtained from the New York City Department of City Planning (NYCDCP[5]). The main zoning categories of interest included commercial, residential, mixed, and park. The areas by zoning category were obtained for each census tract using a Visual Basic for Applications (VBA) program developed in ArcGIS and sequentially the ratios of zoning categories to the total zone area were computed.

The demo-economic data for census tracts was obtained from the 2011 census data provided by U.S. Census Bureau[6] . The main categories of demo-economic data include demographic (e.g., total population, population under 14 and population over 65), economic (e.g., unemployment rate and median income), housing (e.g., median value and household average size), and commuting (e.g., the ratios of commuters by driving alone, carpooling, public transit and walking) data.

All the above data were assembled with the crash data and dangerous driving events to create the final dataset for modeling. The descriptions and descriptive statistics of specific variables derived from the dataset are presented in Table 1.

5. Discussion of results

The proposed MVCAR model was developed to model the crash frequencies in different times of day. For comparison purpose, the Poisson-Gamma (PG) model and univariate conditional autoregressive (UCAR) model were also developed in this study. To obtain the most parsimonious model, stepwise selection method based on Akaike Information Criteria (AIC) (Huang et al., 2009; Yamashita et al., 2007) was employed in selecting explanatory variables using the stepAIC function in R (Ripley et al., 2013) before estimating the final model in the Bayesian framework. Generally speaking, stepwise method usually starts with a model with no input variables. It then selects one significant variable to add to the current model by comparing the criterion value (such as AIC) of all the models with one additional variable. The algorithm stops if no more variables are selected. Please refer to Yamashita et al. (2007) for a detailed discussion on the stepwise selection method. The variables kept by the stepwise AIC method were further examined according to their significance and contribution to goodness-of-fit. Different explanatory variables were selected to model the number of crashes in different time periods. Note that no socio-demographic variables are included in the final model after this

[1] Source: https://opendata.cityofnewyork.us/.

[2] Source: https://gis.ny.gov/gisdata.

[3] Source: http://www.nymtc.org/project/bpm/bpmindex.html.

[4] Source: http://web.mta.info/developers/download.html.

[5] Source: http://www.nyc.gov/html/dcp/html/bytes/dwn_pluto_mappluto.shtml.

[6] Source: http://factfinder.census.gov.

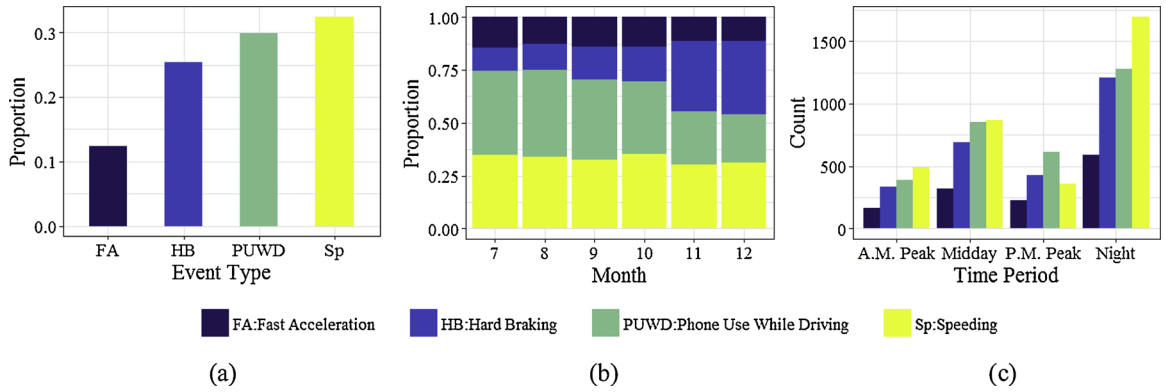

Fig. 2. Distribution of four types of dangerous driving behaviors.

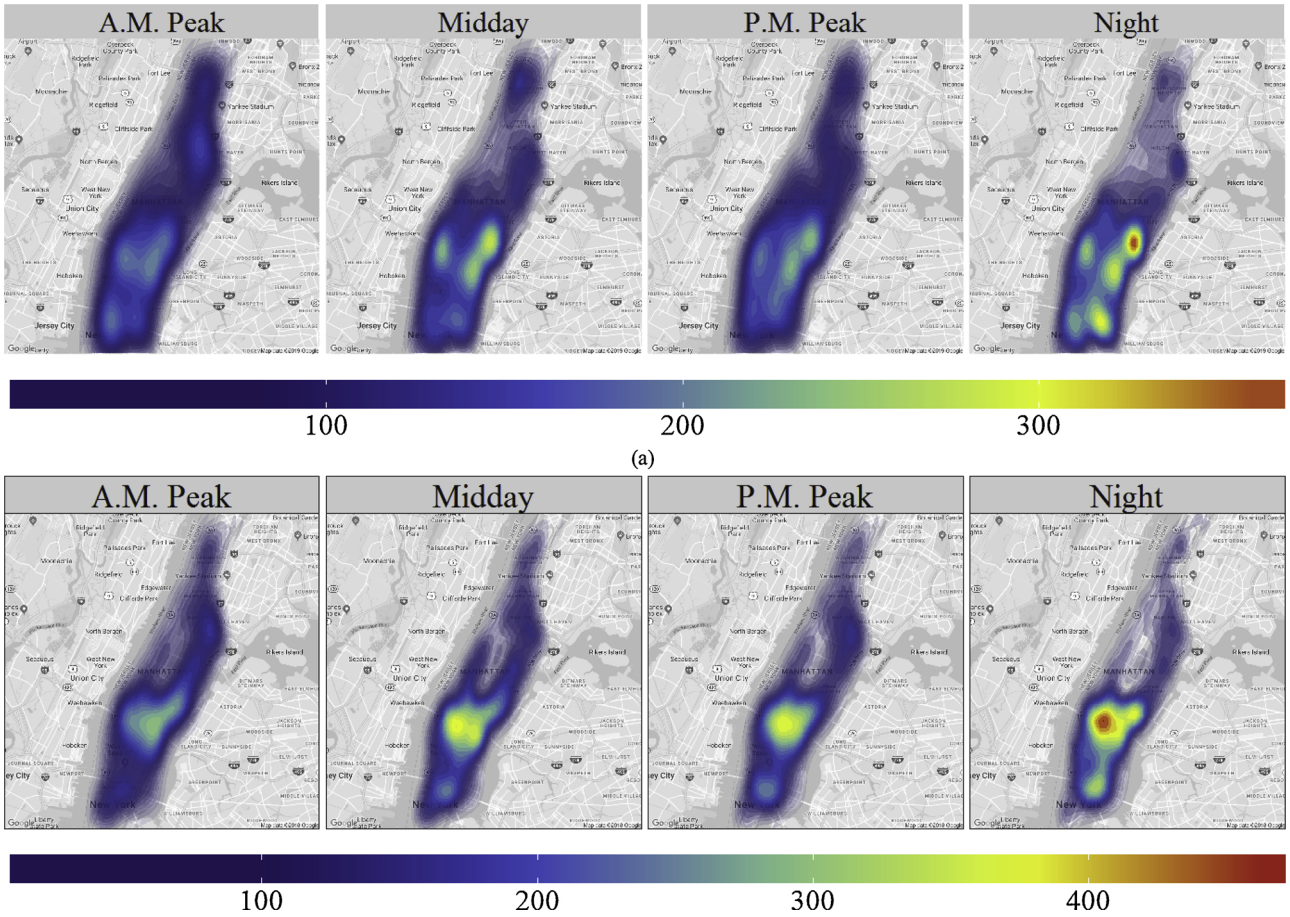

Fig. 3. Spatial distributions in different times of day for (a) dangerous driving events and (b) crashes.

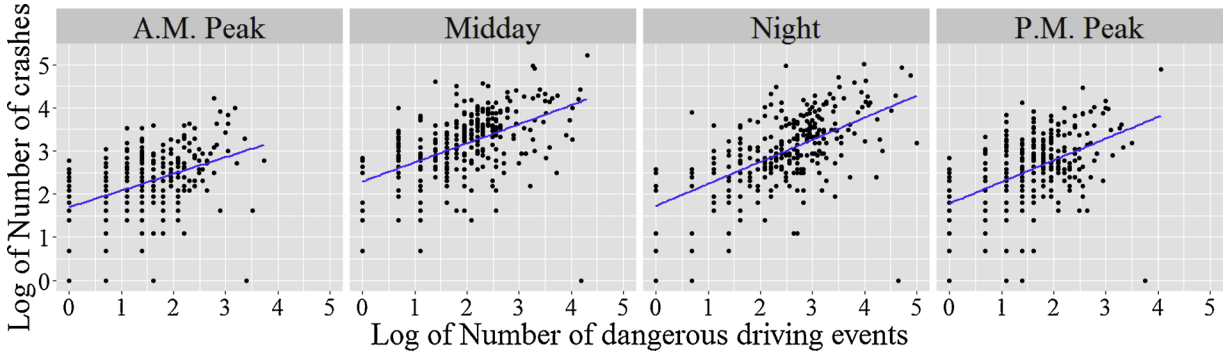

Fig. 4. Scatterplots between number of crashes and number of dangerous driving events for four time periods.

Table 1
Description and descriptive statistics of data (n = 282 census tracts).

Variable	Description	Mean	Standard Deviation
Crash			
AM Peak	Count of crashes in the AM Peak period	4.94	5.55
Midday	Count of crashes in the Midday period	9.69	11.57
PM Peak	Count of crashes in the PM Peak period	5.75	6.29
Night	Count of crashes in the Night period	16.90	20.04
Dangerous driving events			
AM Peak	Count of dangerous driving events in the AM Peak period	11.20	9.05
Midday	Count of dangerous driving events in the Midday period	29.93	23.37
PM Peak	Count of dangerous driving events in the PM Peak period	16.82	14.52
Night	Count of dangerous driving events in the Night period	26.31	23.66
Transportation			
LogVMT	The logarithm of annual average daily vehicle miles traveled (veh.mile)	9.56	1.69
Truck ratio	The average ratio of truck flow to total flow	0.06	0.04
Avenue length	Total length of avenues (mile)	0.39	0.27
Avenue length ratio	The ratio of the total avenue length (mile) and the total road length (mile)	0.37	0.29
Intersection number	Count of intersections in each census tract	11.86	6.9
Subway station number	Count of subway stations in each census tract	0.52	0.92
Bus stop number	Count of bus stops in each census tract	8.09	5.92
Land use			
Commercial ratio	The ratio of commercial zone area to the whole area	0.3	0.33
Residential ratio	The ratio of residential zone area to the whole area	0.56	0.35
Mixed ratio	The ratio of mixed zone area to the whole area	0.07	0.17
Park ratio	The ratio of park area to the whole area	0.06	0.14
Demo-economic			
Population	Total population (10^3)	5.56	3.14
Population under 14	Population under 14 years (10^3)	0.7	0.54
Population over 65	Population 65 years and over (10^3)	0.74	0.56
Median age	Median age of population	37.36	7.55
Median income	Median income per household (10^3 $)	76.41	44.49
Median housing value	Median value of housing (10^3 $)	615.6	235.36
Unemployment rate	Share of the labor force that is unemployed	0.09	0.05

stepwise AIC selection method.

The summary of DIC statistics are presented in Table 2. DIC value of the MVCAR model (6,738) is 317 and 124 less than that of the PG model (7,055) and the UCAR model (6,862), respectively. It suggests that the MVCAR model has the best performance. By including the CAR effect term, the UCAR model outperforms the PG model in light of its lower DIC. Compared to the UCAR model, the MVCAR model which addresses the correlation of different time periods can be regarded as superior since it reduces DIC value by 124 (from 6,862 for the UCAR model to 6,738 for the MVCAR model).

5.1. Temporal effects of risk factors

Regarding the best performance of the MVCAR model, its Bayesian posterior estimates are used for variable interpretation, as presented in Table 3. For comparison purpose, a UCAR model for all-day crashes is estimated and reported in Table 3. The 90% Bayesian Credible Interval (90% BCI) is used to examine the significance of estimations. Estimates can be regarded as significant at the 90% level if the BCIs do not cover 0 and vice versa (Gelman et al., 2014). Excluding intercepts, as shown in Table 3, two parameters are statistically insignificant, namely commercial ratio in night period and subway entrance number in the all-day crash model. Because the lower bounds of the 90% BCI of these two parameters are only slightly less than zero, they were included in the final model specification. As for the dispersion parameter, it is apparent that values are much smaller in the four time periods than in the total crash model. One possible explanation is that unobserved spatial and temporal heterogeneity is accounted for in the MVCAR models for different time periods while only unobserved spatial heterogeneity is accounted for in the UCAR all-day crash model.

Table 2
Summary of DIC values.

Model	DIC
Model 1. Poisson-Gamma (PG) model	7,055
Model 2. Univariate conditional autoregressive (UCAR) model	6,862
Model 3. Multivariate conditional autoregressive (MVCAR) model	6,738

Vehicle miles traveled (VMT) is included in all four time periods and all-day crash models and is found to be positively associated with crash counts. Notice that because of the lack of detailed flow for estimating VMT in each period, potential bias might occur due to using aggregated VMT. A possible solution is to apply hourly factor of AADT to approximate the flow and calculate VMT in each period, which we will not pursue in this study. The increase in the all-day crash counts is the smallest comparing to increases in all the four time periods after controlling for other variables. Effects of VMT in the a.m. peak and p.m. peak periods are similar to each other and relatively higher while midday and night periods are similar to each other and relatively lower. This indicates that relationships between the total exposure variable and crash counts are different at different time periods.

Another variable predictive of crashes is the truck ratio and it shows that higher truck ratio promotes the likelihood of crash counts in all four time periods. Coefficients in a.m. peak and midday periods are similar with each other while coefficients in the p.m. peak, night, and all-day crash model approximate each other. Additionally, avenue length ratio is found to have positive impacts on a.m. peak, midday, and p.m. peak crash counts. After controlling for other variables, effects of avenue length ratio are similar to each other in these three time periods, which indicate relatively stable relationships between avenue length ratio and crash counts.

As for land use patterns, previous studies (Pulugurtha and Sambhara, 2011; Wier et al., 2009) showed that land use patterns could influence the occurrence of crashes. In this study, one noticeable difference between the four time periods and all-day crash model is that

Table 3
Posteriors of the MVCAR models for different times of day and the UCAR all-day crash model.

Variables	AM Peak		Midday		PM Peak		Night		All-day crash model	
	Mean	90% BCI	Mean	90% BCI	Mean	90% BCI	Mean	90% BCI	Mean	90% BCI
Intercept	−1.961	(−2.563, −1.346)	−0.062	(−0.538, 0.345)	−1.700	(−2.378, −1.181)	−0.545	(−1.005, 0.010)	0.821	(0.590, 1.061)
Log(VMT)	0.374	(0.312, 0.435)	0.266	(0.223, 0.315)	0.367	(0.314, 0.434)	0.280	(0.217, 0.328)	0.216	(0.192, 0.240)
Truck ratio	2.743	(1.189, 4.327)	2.722	(1.353, 4.167)	4.243	(2.623, 5.872)	3.934	(2.346, 5.536)	3.734	(2.281, 5.210)
Avenue length ratio	0.249	(0.033, 0.468)	0.238	(0.065, 0.407)	0.210	(0.019, 0.400)	–	–	–	–
Commercial ratio	–	–	–	–	–	–	0.092	(-0.047, 0.234)	0.266	(0.063, 0.464)
Park ratio	–	–	–	–	–	–	–	–	−0.608	(−1.009, −0.210)
Bus stop number	0.024	(0.015, 0.034)	0.026	(0.017, 0.036)	0.033	(0.023, 0.043)	0.031	(0.022, 0.041)	0.023	(0.013, 0.035)
Subway entrance number	–	–	0.009	(0.002, 0.016)	–	–	–	–	0.011	(−0.002, 0.025)
Log(Dangerous driving events count)	0.064	(0.004, 0.123)	0.090	(0.042, 0.138)	0.077	(0.020, 0.134)	0.121	(0.068, 0.174)	0.264	(0.220, 0.320)
Dispersion	0.006	(0.000, 0.017)	0.003	(0.000, 0.010)	0.003	(0.000, 0.010)	0.004	(0.000, 0.012)	0.147	(0.112, 0.184)

the ratio of park area is statistically significant in the all-day crash model while not even selected in the four time periods models. The ratio of commercial area has positive effects on both crash counts during night and all-day crash counts while the ratio of park area has negative effect on the all-day crash count. Positive relationships between crash occurrence and subway entrance number and bus stop number are found in literature (Ukkusuri et al., 2011). In this study, as shown in Table 3, bus stop number variable is included in all four periods and the all-day crash model and is also positively associated with crash counts. Coefficients of bus stop number are similar among a.m. peak, midday, and all-day crash model while coefficients in p.m. peak and night period are similar. Subway entrance number variable is included in both the midday period and the all-day crash model but only significant in the midday period. Effects of subway entrance number variable are similar between the midday period and the all-day crash model after controlling for other variables.

Previous studies (Gitelman et al., 2018; Pande et al., 2017) showed that dangerous driving events could influence the occurrence of crashes. As shown in Table 3, dangerous driving events in models for different time periods and the all-day crash model are found to have statistically significant positive effects on crash occurrence. One-unit increase in the logarithm of the number of dangerous driving events is expected to be associated with an increase in crash counts of 6.6% ($e^{0.064}$-1) in a.m. peak, 9.4% ($e^{0.090}$-1) in midday, 8.0% ($e^{0.077}$-1) in p.m. peak, 12.9% ($e^{0.121}$-1) in night after controlling for other variables. Posterior distributions of coefficients of dangerous driving events in four time periods are shown in Fig. 5. Probability density plots of all the parameters are bell-shaped. The coefficient of dangerous driving behaviors at night is the largest of all time periods, which is almost twice of that in the a.m. peak period. One possible explanation is that, poor visibility during the nighttime allows drivers less time to take evasive actions when exposed to risks caused by dangerous driving events.

To further test whether the differences of percentage changes among different time periods are statistically significant, two-sample Kolmogorov-Smirnov tests (Lopes, 2011) were used to test between each pair of posterior distributions in Fig. 5. Two-sample Kolmogorov-Smirnov test was designed to test whether two samples are drawn from the same distribution. It is a nonparametric test, which doesn't make any assumptions about the distributions of the two samples. The outcomes of two-sample Kolmogorov-Smirnov tests are reported in Table 4. A p-value lower than 0.05 indicates a significant difference between a pair of posterior distributions. As can be seen from Table 4, all p-values are less than 0.05, which means that posterior distributions in different time periods are statistically significantly different from each other. This great variation in the safety effects of dangerous driving behaviors can further justify the necessity to develop time-dependent SPFs.

The estimated conditional standard deviation (SD) of the spatial effects of each individual time period and the within-site conditional correlation of the spatial effects of different time periods in the MVCAR model are shown in Table 5. All the estimates in Table 5 were found to be statistically significant (90% BCI do not cover 0) and the results provided evidence for the multivariate spatial autocorrelation of different time periods. The conditional SDs of the spatial effects of the four time periods are 0.976, 1.069, 1.121, and 1.141 respectively, thus one SD higher in spatial effect is associated with 165% ($e^{0.976}$-1), 191% ($e^{1.069}$-1), 207% ($e^{1.121}$-1), 213% ($e^{1.141}$-1) higher expectation in A.M Peak, Midday, P.M. Peak, and Night crash counts, respectively. The SDs of the spatial effects by time periods are of similar levels, because the spatial effects influence the crash count proportionally. It is an expected outcome that the spatial effect of the Night period has a slightly higher SD than the other three time periods, because generally factors contributing to night time crashes are not included for modeling in this study. Regarding the correlation between crashes in different time periods, all the within-site conditional correlation coefficients of spatial effects are positive and greater than 0.8, suggesting a strong shared geographical pattern of risk among different time periods. Another finding is that correlations between consecutive time periods (i.e. A.M. Peak and Midday, Midday and P.M. Peak, P.M. Peak and Night) are higher than those that are not consecutive. It is likely that unobserved factors contributing to crash occurrences during consecutive time periods are more similar.

5.2. Time-dependent hotspot identification

PSI value for each census tract was calculated using Eq. (4). Because census tracts with PSI values less than zero are regarded as safe, these census tracts are colored with dark blue while census tracts with PSI values greater than zero are colored from blue to red based on the PSI values. Census tracts with PSI values greater than zero are identified as hotspots. The spatial distribution of PSIs from the all-day crash model and four time periods are shown in Fig. 6, respectively.

As shown in Fig. 6, for the all-day crash model, hotspots are clustered mainly in Midtown with some hotspots spread across uptown and one cluster in downtown. Regarding hotspots identified in different time periods, generally, the clustering of hotspots in Midtown and Downtown can still be observed, but there are some noticeable differences in hotspots distributions among these time periods, which are summarized below.

a) Hotspots spread widely in the whole Manhattan in the a.m. peak and midday periods with one cluster in Midtown and one cluster in Downtown, while hotspots are clustered mainly in Midtown and Downtown in the p.m. peak and night periods.

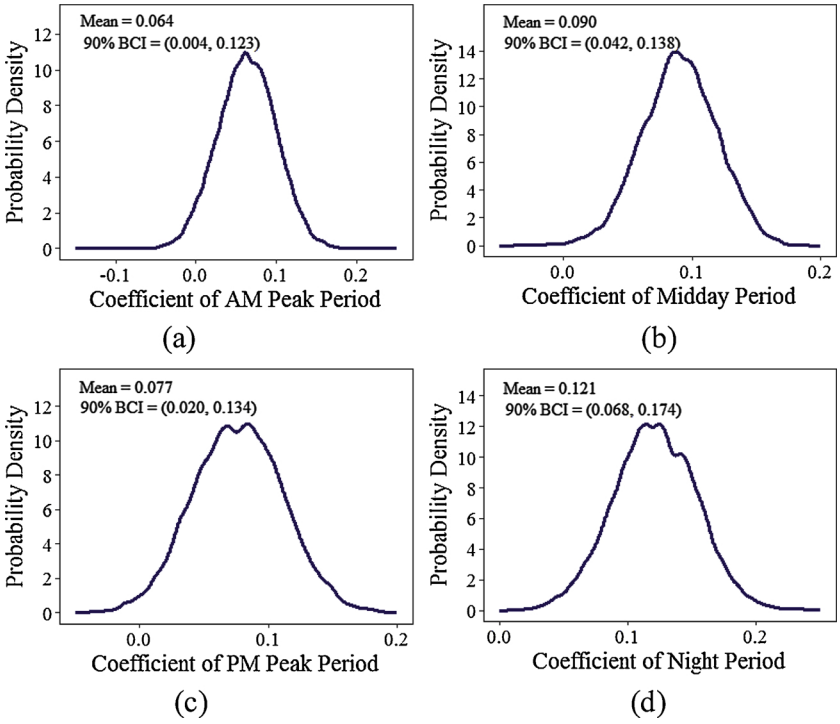

Fig. 5. Posterior distributions of dangerous driving behaviors coefficients in four time periods.

Table 4
Pairwise Kolmogorov-Smirnov test results (p-values) among dangerous driving behaviors posterior distributions.

Time periods	A.M. Peak	Midday	P.M. Peak	Night
A.M. Peak	–	–	–	–
Midday	< 0.001	–	–	–
P.M. Peak	< 0.001	< 0.001	–	–
Night	< 0.001	< 0.001	< 0.001	–

Table 5
Estimates of the conditional standard deviation (SD) of the spatial effects of each individual time period and the within-site conditional correlation of the spatial effects of each pair of time periods.

	Mean	90% BCI
Conditional standard deviation (SD)		
A.M. Peak: σ_{S11}	0.976	(0.877, 1.087)
Midday: σ_{S22}	1.069	(0.979, 1.169)
P.M. Peak: σ_{S33}	1.121	(1.022, 1.230)
Night: σ_{S44}	1.141	(1.044, 1.248)
Within-site conditional correlation		
A.M. Peak vs Midday: $\sigma^2_{S12}/(\sigma_{S11} \times \sigma_{S22})$	0.907	(0.866, 0.938)
A.M. Peak vs P.M. Peak: $\sigma^2_{S13}/(\sigma_{S11} \times \sigma_{S33})$	0.886	(0.835, 0.923)
A.M. Peak vs Night: $\sigma^2_{S14}/(\sigma_{S11} \times \sigma_{S44})$	0.839	(0.774, 0.888)
Midday vs P.M. Peak: $\sigma^2_{S23}/(\sigma_{S22} \times \sigma_{S33})$	0.933	(0.903, 0.954)
Midday vs Night: $\sigma^2_{S24}/(\sigma_{S22} \times \sigma_{S44})$	0.889	(0.849, 0.920)
P.M. Peak vs Night: $\sigma^2_{S34}/(\sigma_{S33} \times \sigma_{S44})$	0.922	(0.890, 0.946)

b) The numbers of identified hotspots in the a.m. peak and the midday periods are higher than those in the rest two time periods. The number of hotspots in the p.m. peak period is the lowest among all the time periods.
c) The census tract with the highest PSI value in the a.m. peak period is different than the census tracts with the highest PSI values in the rest time periods.
d) The number of uniquely identified hotspots in each time period other than the other three are sixteen in the a.m. peak period, nine in the midday period, zero in the p.m. peak period, and eighteen in the night period.

We propose to use the Wilcoxon signed-rank test (Woolson, 2007) to compare the rankings of hotspots identified in different times of day. Similar to the Kolmogorov-Smirnov test, the Wilcoxon signed rank test is also a nonparametric test (Washington et al., 2010). Another reason is that comparing to Wilcoxon rank sum test (also known as Mann-Whitney U test) (Mann and Whitney, 1947; Wilcoxon, 1945), which was designed to test the difference between two independent samples, Wilcoxon signed rank test is useful for comparing two samples for which the observations are paired (Washington et al., 2010). This is appropriate in comparing identified hotpots since PSI values are paired through census tracts. Similarly, because the sample size in our study is the number of census tracts in Manhattan, which is 282 ($\geq$ 25), the large sample normal approximation is also applied here. The outcomes of Wilcoxon signed-rank tests are reported in Table 6. It is found that hotspots identified from all-day crashes are different from those of the p.m. peak at a significance level of 0.05 and different from those of midday and night at a significance level of 0.1. P-values of pairwise tests between the a.m. peak and the midday, the p.m. peak, and the night periods are all less than 0.1 except that between midday and night, indicating significant differences in hotspots identified in different times of day at a significance level of 0.1. The fact that p-values of Wilcoxon signed-rank tests are mostly small justifies the need to identify time-dependent hotspots instead of identifying hotspots aggregately.

5.3. Applications

5.3.1. Developing time-dependent safety countermeasures

The difference among time-dependent hotspots gives us opportunities to develop time-related safety countermeasures that target at each time period individually. One of the countermeasures is to develop time-dependent police patrol plans according to time-dependent hotspots. The concept of developing police patrol plan has been explored in

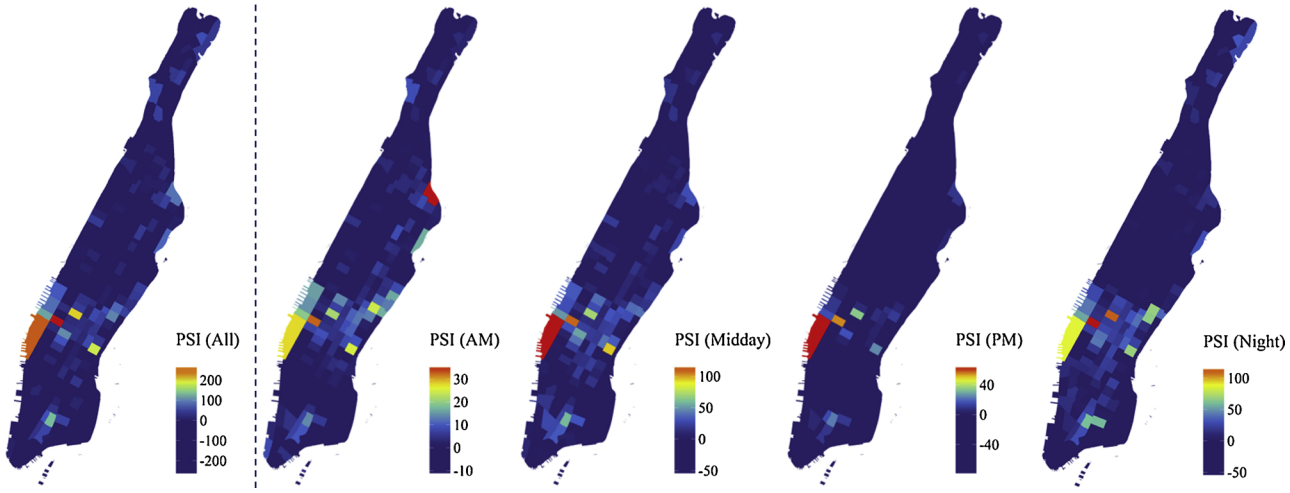

Fig. 6. Hotspots identified using PSI for aggregated UCAR model.

Table 6
Pairwise Wilcoxon signed-rank test results (p-values) of PSI values among four time periods and total crash model.

Time periods	All-day crash model	A.M. Peak	Midday	P.M. Peak	Night
All-day crash model	–	–	–	–	–
A.M. Peak	0.224	–	–	–	–
Midday	0.090*	0.002**	–	–	–
P.M. Peak	< 0.001**	< 0.001**	< 0.001**	–	–
Night	0.084*	0.081*	0.491	< 0.001**	–

(Note: *significant at a level of 0.1, ** significant at a level of 0.05).

literature both in addressing crime hotspots (Camacho-Collados and Liberatore, 2015) and traffic crash hotspots (Kuo et al., 2013). This study can support the idea of developing time-dependent police patrol plans since time-dependent hotspots have been identified and the spatial distributions of hotspots in different time periods are different according to statistical tests (Table 6). It has been shown in the literature (Bagloee and Asadi, 2016) that police presence at busy intersections during busy night outs and weekends highly improves the safety of pedestrians. For our study, for example, the spatial distribution of hotspots in the Night period is more scattered than that in the P.M. Peak period. So, one possible police patrol plan from the P.M. Peak period to the Night period for Manhattan is to add more police force or spread the current police force to cover more areas. Future directions may include developing detailed time-dependent police patrol plans considering police precinct locations, patrolling budget and so on.

Further, other time-dependent countermeasures may be developed by policy makers and transportation engineers. Improving lighting condition has been proved to improve the safety of night times (Sasidharan and Donnell, 2013). So, one possible countermeasure is to consider building lights at locations identified as night time hotspots. In terms of New York City, there are other two application cases that may be of interest. In 2009 and 2010, New York City conducted an Off-Hour Truck Delivery Pilot program to shift a portion of delivery trucks from the regular daytime hours, that is, between 6 a.m. and 7 p.m., to the nighttime off-hours, that is, between 7 p.m. and 6 a.m. (Xie et al., 2015). This truck volume shift may bring more truck involved crashes from day time to night time. So, for those eighteen hotspots that are uniquely identified at night period, countermeasures that can reduce truck involved crashes can be implemented, such as providing exclusive parking spaces for trucks in the central business district (i.e. Midtown Manhattan) at night. Speaking of parking, the second application case in New York City involves parking regulations. Some places in New York City have parking restrictions at different time periods of day. This may impose illegal and unsafe parking that may eventually lead to traffic crashes. Thus, more strict countermeasures may be implemented based on hotspots distribution at different time periods of day in terms of parking. Besides, adjusting signal timing for certain time periods to reduce potential crashes may be another good option for consideration.

5.3.2. *Assessing effects of time-dependent safety countermeasures*

After implementing safety countermeasures, it is important to evaluate their safety effects. Traditionally, crashes over several years both before and after the implementation of safety countermeasures are needed (Hauer, 1997). However, the time-dependent SPF developed in this study and the use of dangerous driving events data may assist in quickly assessing effects of safety countermeasures.

As shown in Table 3, eight variables used in developing the time-dependent SPF are average VMT in 2015, truck ratio, avenue length ratio, commercial ratio, park ratio, bus stop number, subway entrance number, and dangerous driving events count. Notice that the first seven variables are static, meaning that they will not change in the year of 2015, while the dangerous driving events count is dynamic, which can account for the change of traffic conditions, including the implementation of safety countermeasures. So, to assess the effect of a certain safety countermeasure, we could count the number of dangerous driving events after the implementation of the countermeasure in a relatively short period of time and use the developed time-dependent SPF to predict the crash reduction ratio. Note that the count of dangerous driving events may exhibit the regression-to-the-mean (RTM) phenomenon. However, when the number of dangerous driving events is relatively large as it is likely the case in the future, the RTM phenomenon may not be noticeable anymore. Nonetheless, further examination is needed before using the count of dangerous driving events to conduct before after analysis.

For example, more police forces have been relocated to one hotspot in the night period. If, by observation, we know there is a 2 unit decrease in the logarithm of the dangerous driving events count in this census tract immediately after adding more police force, then we can predict a decrease of crash count by 21.5% ($e^{0.121\times(-2)} - 1$) based on the estimated coefficient of dangerous driving events at night (shown in Table 3). The dynamic nature of dangerous driving events and the time-dependent SPF developed in this study have the potential to give transportation practitioners a quick estimate of effects of safety countermeasures.

6. Summary and conclusions

This study examined time-dependent safety performance by leveraging dangerous driving event data. Dangerous driving events were directly related with crash occurrence in different time periods. The multivariate conditional autoregressive (MVCAR) model, which can jointly account for the spatial and temporal dependences of crash observations, was found to achieve the best performance in associating risk factors with crash occurrence. Variation in estimated coefficients in models for different time periods and the whole day have been noticed.

Results of two-sample Kolmogorov-Smirnov tests show that the posterior distributions of coefficients of dangerous driving events at different times of day are statistically significantly different from each other. It highlights the necessity to develop time-dependent SPFs to account for the varying temporal effect of dangerous driving events on safety. Time-dependent hotspots were identified using Potential for Safety Improvement (PSI) metric, which has been widely used for hotspot ranking. The Wilcoxon signed-rank test was used to test the difference between hotspot ranking for different time periods and the ranking for all-day crashes. Through pairwise comparisons, results show that hotspots of the all-day crash model are different from those of the p.m. peak at a significance level of 0.05 and different from those of midday and night at a significance level of 0.1. Hotspots identified by times of day are found to be mostly different from each other.

As discussed in 5.2, the number of uniquely identified hotspots in each time period is different from one to another, with night as the highest. Uniquely identified hotspots suggest that the geometric design of these census tracts may not cause the high frequency of crashes but rather time-dependent risk factors. This can be of interest to traffic authorities and policy makers in reinforcing time-dependent traffic calming measures, especially the practical ones like police patrols, signal-timing adjustment, and so on. Note that even though we have high fidelity dangerous driving event data, we don't want to completely abandon the concept of SPF since it is well-established and widely accepted. We improved SPFs and SPF-based safety implementations like hotspot identification leveraging the high fidelity data. This study also shows a promising future of combining dangerous driving event data collected by smartphones with crashes to provide better understanding of safety management. In this study we improved our current understanding of SPF by investigating the dependency among crash counts across time periods. This is made possible by including the dangerous driving event data. With the same logic of treating high fidelity dangerous driving events data as exposure for crashes, one interesting future topic is to include dangerous driving events data for real-time crash prediction and evaluate the improvement in the prediction accuracy. Also, we could extend the idea of combining dangerous driving event data and crashes together to other aspect of traffic safety management, such as before-after analysis for safety-related countermeasures (Xie et al., 2018a, 2018b). Further, driver information can be integrated into the analysis to account for the dependence between dangerous driving events generated from the various types of drivers in the future. Real world connected vehicle data might be used to extract more crash surrogate measures and dangerous driving events (Xie et al., 2018b). Detailed analysis of funding allocation regarding time dependent safety countermeasures can be performed (Melachrinoudis & Kozanidis, 2002).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The work is partially funded by Connected Cities for Smart Mobility towards Accessible and Resilient Transportation (C2SMART) Center at New York University (NYU). The authors would like to thank the New York State Department of Transportation, New York City Department of City Planning, Metropolitan Transportation Authority, New York Metropolitan Transportation Council and U.S. Census Bureau for publicly providing data used in this study. The authors would also like to thank Zendrive Inc for providing dangerous driving event data used in this study. The contents of this paper present views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents of the paper do not reflect the official views or policies of the agencies.

References

Abdel-Aty, M., Abdalla, F., 2004. Linking roadway geometrics and real-time traffic characteristics to model daytime freeway crashes: generalized estimating equations for correlated data. Transp. Res. Rec.: J. Transp. Res. Board (1897), 106–115.
Abdel-Aty, M., Lee, J., Siddiqui, C., Choi, K., 2013. Geographical unit based analysis in the context of transportation safety planning. Transp. Res. Part A Policy Pract. 49, 62–75.
Aguero-Valverde, J., 2013a. Full Bayes Poisson gamma, poisson lognormal, and zero inflated random effects models: comparing the precision of crash frequency estimates. Accid. Anal. Prev. 50, 289–297.
Aguero-Valverde, J., 2013b. Multivariate spatial models of excess crash frequency at area level: case of Costa Rica. Accid. Anal. Prev. 59, 365–373.
Allen, B.L., Shin, B.T., Cooper, P.J., 1978. Analysis of traffic conflicts and collisions. Transp. Res. Rec.: J. Transp. Res. Board (667), 67–74.
Bagloee, S.A., Asadi, M., 2016. Crash analysis at intersections in the CBD: a survival analysis model. Transp. Res. Part A Policy Pract. 94, 558–572.
Balagh, A.K.G., Naderkhani, F., Makis, V., 2014. Highway accident modeling and forecasting in winter. Transp. Res. Part A Policy Pract. 59, 384–396.
Baltusis, P., 2004. On Board Vehicle Diagnostics. Retrieved from. .
Broughton, P., Fuller, R., Stradling, S., Gormley, M., Kinnear, N., O'dolan, C., Hannigan, B., 2009. Conditions for speeding behaviour: a comparison of car drivers and powered two wheeled riders. Transp. Res. Part F Traffic Psychol. Behav. 12 (5), 417–427.
Camacho-Collados, M., Liberatore, F., 2015. A decision support system for predictive police patrolling. Decis. Support Syst. 75, 25–37.
Cooper, D.F., Ferguson, N., 1976. Traffic studies at T-Junctions. 2. A conflict simulation record. Traffic Eng. Control 17 Analytic.
De Pelsmacker, P., Janssens, W., 2007. The effect of norms, attitudes and habits on speeding behavior: scale development and model building and estimation. Accid. Anal. Prev. 39 (1), 6–15.
Delhaye, E., 2006. Traffic safety: speed limits, strict liability and a km tax. Transp. Res. Part A Policy Pract. 40 (3), 205–226.
Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., et al., 2006a. The 100-Car Naturalistic Driving Study, Phase II-results of the 100-Car Field Experiment. Retrieved from. .
Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., et al., 2006b. The 100-Car Naturalistic Driving Study. Phase 2: Results of the 100-Car Field Experiment. Retrieved from. .
Doecke, S., Grant, A., Anderson, R.W., 2015. The real-world safety potential of connected vehicle technology. Traffic Inj. Prev. 16 (sup1), S31–S35.
Dunlop, M.D., Roper, M., Elliot, M., McCartan, R., McGregor, B., 2016. Using smartphones in cities to crowdsource dangerous road sections and give effective in-car warnings. Paper Presented at the Proceedings of the SEACHI 2016 on Smart Cities for Better Living with HCI and UX.
Elliott, M.A., Armitage, C.J., Baughan, C.J., 2003. Drivers' compliance with speed limits: an application of the theory of planned behavior. J. Appl. Psychol. 88 (5), 964.
Elyasi, M.R., Saffarzade, M., Boroujerdian, A.M., 2016. A novel dynamic segmentation model for identification and prioritization of black spots based on the pattern of potential for safety improvement. Transp. Res. Part A Policy Pract. 91, 346–357.
Fazeen, M., Gozick, B., Dantu, R., Bhukhiya, M., González, M.C., 2012. Safe driving using mobile phones. IEEE Trans. Intell. Transp. Syst. 13 (3), 1462–1468.
Feng, Y., Head, K.L., Khoshmagham, S., Zamanipour, M., 2015. A real-time adaptive signal control in a connected vehicle environment. Transp. Res. Part C Emerg. Technol. 55, 460–473.
Folkard, S., 1997. Black times: temporal determinants of transport safety. Accid. Anal. Prev. 29 (4), 417–430.
Galgano, S., Talas, M., Opie, K., Marsico, M., Weeks, A., Wang, Y., et al., 2016. Connected Vehicle Pilot Deployment Program Phase 1: Performance Measurement and Evaluation Support Plan: New York City. Retrieved from. .
Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2014. Bayesian Data Analysis Vol. 2 CRC press, Boca Raton, FL.
Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 457–472.
Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1998. Markov Chain Monte Carlo in Practice. Chapman & Hall, Boca Raton, Florida.
Gitelman, V., Bekhor, S., Doveh, E., Pesahov, F., Carmel, R., Morik, S., 2018. Exploring relationships between driving events identified by in-vehicle data recorders, infrastructure characteristics and road crashes. Transp. Res. Part C Emerg. Technol. 91, 156–175.

Golob, T.F., Recker, W.W., 2004. A method for relating type of crash to traffic flow characteristics on urban freeways. Transp. Res. Part A Policy Pract. 38 (1), 53–80.
Guo, F., Fang, Y., 2013. Individual driver risk assessment using naturalistic driving data. Accid. Anal. Prev. 61, 3–9.
Guo, F., Klauer, S.G., Hankey, J.M., Dingus, T.A., 2010. Near crashes as crash surrogate for naturalistic driving studies. Transp. Res. Rec. 2147 (1), 66–74.
Hauer, E., 1995. On exposure and accident rate. Traffic Eng. Control 36 (3).
Hauer, E., 1997. Observational before/after Studies in Road Safety. Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety.
Hauer, E., Kononov, J., Allery, B., Griffith, M., 2002. Screening the road network for sites with promise. Transp. Res. Rec.: J. Transp. Res. Board (1784), 27–32.
Hayward, J.C., 1972. Near miss determination through use of a scale of danger. Transp. Res. Rec.: J. Transp. Res. Board 384, 24–34.
Howe, G., Xu, G., Hoover, D., Elsasser, D., Barickman, F., 2016. Commercial Connected Vehicle Test Procedure Development and Test Results–Emergency Electronic Brake Light. Retrieved from. .
Huang, H., Chin, H., Haque, M., 2009. Empirical evaluation of alternative approaches in identifying crash hot spots: naive ranking, empirical Bayes, and full Bayes methods. Transp. Res. Rec.: J. Transp. Res. Board (2103), 32–41.
Ivan, J.N., Wang, C., Bernardo, N.R., 2000. Explaining two-lane highway crash rates using land use and hourly exposure. Accid. Anal. Prev. 32 (6), 787–795.
Johnson, D.A., Trivedi, M.M., 2011. Driving style recognition using a smartphone as a sensor platform. Paper Presented at the Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on.
Kahle, D., Wickham, H., 2013. ggmap: Spatial Visualization with ggplot2. R J. 5 (1).
Kanhere, S.S., 2011. Participatory sensing: crowdsourcing data from mobile smartphones in urban spaces. In: Paper Presented at the Mobile Data Management (MDM). 2011 12th IEEE International Conference on.
Klauer, S.G., Guo, F., Simons-Morton, B.G., Ouimet, M.C., Lee, S.E., Dingus, T.A., 2014. Distracted driving and risk of road crashes among novice and experienced drivers. N. Engl. J. Med. 370 (1), 54–59.
Kuo, P.-F., Lord, D., Walden, T.D., 2013. Using geographical information systems to organize police patrol routes effectively by grouping hotspots of crash and crime data. J. Transp. Geogr. 30, 138–148.
Lan, B., Persaud, B., Lyon, C., Bhim, R., 2009. Validation of a full Bayes methodology for observational before-after road safety studies and application to evaluation of rural signal conversions. Accid. Anal. Prev. 41 (3), 574–580. https://doi.org/10.1016/j.aap.2009.02.010.
Lee, J., Brown, T., Fiorentino, D., Fell, J.C., Traube, E., Nadler, E., 2011. Using Vehicle-based Sensors of Driver Behavior to Detect Alcohol Impairment.
Lee, J.D., 2008. Fifty years of driving safety research. Hum. Factors 50 (3), 521–528.
Lenné, M.G., Triggs, T.J., Redman, J.R., 1997. Time of day variations in driving performance. Accid. Anal. Prev. 29 (4), 431–437.
Lopes, R.H., 2011. Kolmogorov-smirnov Test International Encyclopedia of Statistical Science. Springer, pp. 718–720.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transp. Res. Part A Policy Pract. 44 (5), 291–305.
Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. 50–60.
Marquis, R., Wang, X., 2015. Investigating temporal effects on truck accident occurrences in Manhattan, New York City. Transp. Res. Rec.: J. Transp. Res. Board (2517), 10–17.
Martin, J.-L., 2002. Relationship between crash rate and hourly traffic flow on interurban motorways. Accid. Anal. Prev. 34 (5), 619–629.
Maze, T.H., Agarwal, M., Burchett, G., 2006. Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. Transp. Res. Rec. 1948 (1), 170–176.
Melachrinoudis, E., Kozanidis, G., 2002. A mixed integer knapsack model for allocating funds to highway safety improvements. Transp. Res. Part A Policy Pract. 36 (9), 789–803.
Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. Accid. Anal. Prev. 39 (3), 459–468. https://doi.org/10.1016/j.aap.2006.08.002.
Pahukula, J., Hernandez, S., Unnikrishnan, A., 2015. A time of day analysis of crashes involving large trucks in urban areas. Accid. Anal. Prev. 75, 155–163.
Paleti, R., Eluru, N., Bhat, C.R., 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. Accid. Anal. Prev. 42 (6), 1839–1854.
Pande, A., Chand, S., Saxena, N., Dixit, V., Loy, J., Wolshon, B., Kent, J.D., 2017. A preliminary investigation of the relationships between historical crash and naturalistic driving. Accid. Anal. Prev. 101, 107–116.
Persaud, B., Lyon, C., Nguyen, T., 1999. Empirical Bayes procedure for ranking sites for safety investigation by potential for safety improvement. Transp. Res. Rec.: J. Transp. Res. Board (1665), 7–12.
Plainis, S., Murray, I., 2002. Reaction times as an index of visual conspicuity when driving at night. Ophthalmic Physiol. Opt. 22 (5), 409–415.
Pulugurtha, S.S., Sambhara, V.R., 2011. Pedestrian crash estimation models for signalized intersections. Accid. Anal. Prev. 43 (1), 439–446.
Qin, X., Ivan, J.N., Ravishanker, N., Liu, J., Tepas, D., 2006. Bayesian estimation of hourly exposure functions by crash type and time of day. Accid. Anal. Prev. 38 (6), 1071–1080.
Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accid. Anal. Prev. 40 (4), 1486–1497.
Redelmeier, D.A., Tibshirani, R.J., 1997. Association between cellular-telephone calls and motor vehicle collisions. N. Engl. J. Med. 336 (7), 453–458.
Richard, C.M., Brown, J.L., Atkins, R., Divekar, G., 2017. Using naturalistic driving data to develop a typology of speeding episodes. Transp. Res. Rec. 2659 (1), 91–97.
Richard, C.M., Campbell, J.L., Lichty, M.G., Brown, J.L., Chrysler, S., Lee, J.D., et al., 2012. Motivations for Speeding, Volume I: Summary Report. Retrieved from. National Highway Traffic Safety Administration.
Ripley, B., Venables, B., Bates, D.M., Hornik, K., Gebhardt, A., Firth, D., Ripley, M.B., 2013. Package 'mass'. Cran R.
Ryder, B., Dahlinger, A., Gahr, B., Zundritsch, P., Wortmann, F., Fleisch, E., 2018. Spatial prediction of traffic accidents with critical driving events–insights from a nationwide field study. Transp. Res. Part A Policy Pract.
Sasidharan, L., Donnell, E.T., 2013. Application of propensity scores and potential outcomes to estimate effectiveness of traffic safety countermeasures: Exploratory analysis using intersection lighting data. Accid. Anal. Prev. 50, 539–553.
Sinharay, S., 2003. Assessing convergence of the Markov chain Monte Carlo algorithms: a review. ETS Res. Rep. Ser. 2003 (1).
Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2003. WinBUGS User Manual: Version.
Spiegelhalter, D.J., Best, N.G., Carlin, B.R., van der Linde, A., 2002a. Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B-Stat. Methodol. 64, 583–616. https://doi.org/10.1111/1467-9868.00353.
Spiegelhalter, D.J., Best, N.G., Carlin, B.R., van der Linde, A., 2002b. Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B Methodol. 64, 583–616. https://doi.org/10.1111/1467-9868.00353.
Stipancic, J., Miranda-Moreno, L., Saunier, N., 2018. Vehicle manoeuvers as surrogate safety measures: Extracting data from the gps-enabled smartphones of regular drivers. Accid. Anal. Prev. 115, 160–169.
Sutton, H., 2017. Study finds unsafe driving practices run rampant near campuses. Campus Secur. Rep. 14 (8) 9-9.
Thomas, A., Best, N., Lunn, D., Arnold, R., Spiegelhalter, D., 2004. GeoBugs User Manual. Medical Research Council Biostatistics Unit, Cambridge.
Toledo, T., Musicant, O., Lotan, T., 2008. In-vehicle data recorders for monitoring and feedback on drivers' behavior. Transp. Res. Part C Emerg. Technol. 16 (3), 320–331.
Ukkusuri, S., Hasan, S., Aziz, H., 2011. Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency. Transp. Res. Rec.: J. Transp. Res. Board (2237), 98–106.
Wang, X., Xie, K., Abdel-Aty, M., Chen, X., Tremont, P.J., 2014. Systematic approach to hazardous-intersection identification and countermeasure development. J. Transp. Eng. 140 (6), 04014022.
Washington, S.P., Karlaftis, M.G., Mannering, F., 2010. Statistical and Econometric Methods for Transportation Data Analysis. CRC press.
White, K.M., Hyde, M.K., Walsh, S.P., Watson, B., 2010. Mobile phone use while driving: an investigation of the beliefs influencing drivers' hands-free and hand-held mobile phone use. Transp. Res. Part F Traffic Psychol. Behav. 13 (1), 9–20.
White, M.P., Eiser, J.R., Harris, P.R., 2004. Risk perceptions of mobile phone use while driving. Risk Anal. 24 (2), 323–334.
Wier, M., Weintraub, J., Humphreys, E.H., Seto, E., Bhatia, R., 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accid. Anal. Prev. 41 (1), 137–145.
Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biom. Bull. 1 (6), 80–83.
Woolson, R., 2007. Wilcoxon signed-Rank Test. Wiley Encyclopedia of Clinical Trials. pp. 1–3.
Wu, K., Jovanis, P.P., 2011. Defining, screening, and validating crash surrogate events using naturalistic driving data. Paper Presented at the Proceedings of the 3rd International Conference on Road Safety and Simulation.
Xie, K., Ozbay, K., Kurkcu, A., Yang, H., 2017. Analysis of traffic crashes involving pedestrians using big data: investigation of contributing factors and identification of hotspots. Risk Anal. 1–18. https://doi.org/10.1111/risa.12785.
Xie, K., Ozbay, K., Yang, H., Holguín-Veras, J., Morgul, E.F., 2015. Modeling safety impacts of off-hour delivery programs in urban areas. Transp. Res. Rec.: J. Transp. Res. Board (2478), 19–27.
Xie, K., Ozbay, K., Yang, H., Yang, D., 2018a. A new methodology for before–After safety assessment using survival analysis and longitudinal data. Risk Anal.
Xie, K., Wang, X., Huang, H., Chen, X., 2013. Corridor-level signalized intersection safety analysis in Shanghai, China using Bayesian hierarchical models. Accid. Anal. Prev. 50, 25–33. https://doi.org/10.1016/j.aap.2012.10.003.
Xie, K., Yang, D., Ozbay, K., Yang, H., 2018b. Use of real-world connected vehicle data in identifying high-risk locations based on a new surrogate safety measure. Accid. Anal. Prev.
Yamashita, T., Yamashita, K., Kamimura, R., 2007. A stepwise AIC method for variable selection in linear regression. Commun. Stat.—Theory Methods 36 (13), 2395–2403.
Zendrive, 2018. Zendrive. Retrieved from. https://www.zendrive.com/.
Zou, W., Wang, X., Wang, Y., 2016. A Multivariate Spatial-time of Day Analysis of Truck Crash Frequency Across Neighborhoods in New York City Spatial Econometrics: Qualitative and Limited Dependent Variables. Emerald Group Publishing Limited, pp. 195–219.