Accident Analysis and Prevention 132 (2019) 105249

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence

Huiying Wen[a,b], Xuan Zhang[a,b], Qiang Zeng[a,b,c,*], N.N. Sze[d]

[a] School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, PR China
[b] Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University Road #2, Nanjing 211189, PR China
[c] Department of Electrical Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, PR China
[d] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, PR China

A R T I C L E I N F O

Keywords:
Freeway safety
Roadway alignment
Weather condition
Interaction effect
Spatio-Temporal model

A B S T R A C T

This study attempts to examine the main and interaction effects of roadway and weather conditions on crash incidence, using the comprehensive crash, traffic and weather data from the Kaiyang Freeway in China in 2014. The dependent variable is monthly crash count on a roadway segment (with homogeneous horizontal and vertical profiles). A Bayesian spatio-temporal model is proposed to measure the association between crash frequency and possible risk factors including traffic composition, presence of curve and slope, weather conditions, and their interactions. The proposed model can also accommodate the unstructured random effect, and spatio-temporal correlation and interactions. Results of parameter estimation indicate that the interactions between wind speed and slope, between precipitation and curve, and between visibility and slope are significantly correlated to the increase in the freeway crash risk, while the interaction between precipitation and slope is significantly correlated to the reduction in the freeway crash risk, respectively. These findings are indicative of the design and implementation of real-time traffic management and control measures, e.g. variable message sign, that could mitigate the crash risk under the adverse weather conditions.

1. Introduction

In the past decade, freeway safety issues have been of increasing concerns of transport authorities and researchers around the world (Ahmed et al., 2011; De Luca and Dell'Acqua, 2012; Jonkers et al., 2011; Sarhan et al., 2008; Zeng et al., 2017a). In particular, the total road length and annual total vehicle mileage of freeways are very high in the countries like the United States (Gaweesh et al., 2019; Shew et al., 2013; Ye et al., 2013; Yu and Abdel-Aty, 2013; Yu et al., 2013) and China (Hou et al., 2018a, b, 2019; Lyu et al., 2018; Ma et al., 2017b; Wen et al., 2018, 2019). Although the crash rates (crash per vehicle mileage) on freeways might be lower than that of other roadway types, e.g. urban roads (Ahmed et al., 2011), the propensity of mortality and severe injury of crashes on freeways could be higher than that of their counterparts, because of the higher vehicular speed and proportion of heavy vehicles on freeways (Zeng et al., 2017a). Taking China as an example, freeway crashes accounted for 5% of overall road crashes only in 2015, however, 10% of road fatalities were attributed to freeways (Zeng et al., 2019). This could indeed have significant impacts on the economic burden and well-being of the society (Hou et al., 2018b).

Because of the significant social implications, it is of great essence to have a good understanding of the factors contributing to the risks of crash and injury on the freeways, and therefore, effective safety countermeasures could be implemented (Lord and Mannering, 2010). In the past two decades, numerous researchers have attempted to develop crash prediction models to measure the relationship between crash frequency and possible contributory factors, including traffic characteristics (e.g., traffic volume and composition, and average speed) (Zeng et al., 2017a; Hou et al., 2018b; Yu et al., 2013), roadway infrastructure (e.g., segment length, number of lanes, horizontal and vertical alignment, and shoulder width) (Ahmed et al., 2011; Gaweesh et al., 2019; Wen et al., 2019), and weather conditions (e.g., precipitation, visibility, temperature, and wind speed) (Gaweesh et al., 2019; Hou et al., 2018b; Yu et al., 2013), based on the historical crash records. However, the majority of the prediction models for freeway crash risk are highly aggregated, i.e. the observation unit is year or season. This may result in possible information loss and bias in parameter estimation, especially for time-varying factors (Lord and Mannering, 2010; Zeng et al., 2017a), such as traffic flow, traffic

* Corresponding author at: School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, PR China.
E-mail addresses: hywen@scut.edu.cn (H. Wen), zhangxuanscut@163.com (X. Zhang), zengqiang@scut.edu.cn (Q. Zeng), tony.nn.sze@polyu.edu.hk (N.N. Sze).

https://doi.org/10.1016/j.aap.2019.07.025
Received 12 March 2019; Received in revised form 17 June 2019; Accepted 25 July 2019
Available online 12 August 2019
0001-4575/

composition and weather conditions.

Another drawback of existing freeway safety studies is that they assume the effect of an explanatory variable to be homogeneous across different subgroups of observation. However, the effect of road geometry on freeway crash risk can be modified by the weather condition, given the same traffic condition. For instances, the impact of vertical grade on crash frequency is partially attributed to the varied sight distance (Labi, 2011), which is also associated with visibility (Yu et al., 2013). Such modification of the effect of an explanatory variable on the outcome, by a third variable, is known as interaction. Indeed, the interaction effects reflect the cumulative effect of two or more explanatory variables, which are not acting independently, on the crash risk. It could provide insight into the effective traffic management and control strategies that could enhance the safety performance of the freeway (Jung et al., 2014; Shankar et al., 2004). In particular, Shankar et al. (2004) attempted the relationship between possible contributory factors and roadside crashes, based on the one-year crash data of 381 one-mile long roadway segments (including freeways, principal, minor, and collector arterials) in the State of Washington. Results of a zero-inflated negative binomial (NB) model indicated that presence of intersection and increase in snow depth could modify the association between roadside crash frequency and horizontal curvature. Jung et al. (2014) applied the traditional NB regression approach to evaluate the effects of road design and traffic management & control attributes on two-vehicle crashes under rainy condition, based on the crash data of selected freeways in the State of Wisconsin. Results indicated that there were significant interactions between speed limit and presence of off-ramp, and number of lanes and pavement type on the crash risk. Significant interaction effects were also revealed in other studies on the crash injury severity (Ahmed et al., 2018; Ponnaluri, 2016; Zhai et al., 2019) and real-time crash risk (Wu et al., 2018) on freeways. Indeed, weather conditions can modify the driver behavior, and therefore the association between possible risk factors and crash occurrence on the freeway. The accumulated effects of hazardous roadway attributes on safety risk might be different under different weather conditions, i.e., rainfall, temperature, and wind speed. To the best of our knowledge, there is no reported research which focuses on the interactions between weather condition and possible roadway characteristics on the crash risk of freeway.

From the methodological perspective, the conventional NB approach (with and without interaction effects) could have deficiency when modeling the crash occurrence based on panel data. In particular, the conventional NB approach could not account for the effects of unobserved heterogeneity and spatio-temporal correlation (Lord and Mannering, 2010; Mannering and Bhat, 2014). To address the issues of unobserved heterogeneity, advanced modeling approach including random-parameters (Hou et al., 2018b), Markov switching (Malyshkina et al., 2009), and latent class/finite mixture structure (Zou et al., 2013) have been proposed. Meanwhile, some studies suggested that unobserved heterogeneity could be attributed to the incorrect functional form (Mannering et al., 2016), missing interactions (Wood et al., 2016), and spatial and temporal correlations (Barua et al., 2016; Cheng et al., 2018b). Recently, Bayesian spatial and temporal approaches have been proposed to take into account the effect of unobserved heterogeneity and spatio-temporal correlation for the crash prediction (Aguero-Valverde and Jovanis, 2006, 2008; Cheng et al., 2018a, 2018b; Dong et al., 2016; Huang et al., 2016; Ma et al., 2017a; Zeng and Huang, 2014; Zeng et al., 2017a, 2017b, 2018). In this study, we attempt to apply a Bayesian spatio-temporal approach to model the crash occurrence on freeway, using the time-series crash data on consecutive freeway segments. The observation unit of current study is one-month crash incidence on consecutive road segments over a long freeway sketch. Therefore, the effect of the monthly variations in traffic and weather conditions on the crash risk can be captured. It is expected that the spatio-temporal correlation between observations should be prevalent. In addition, the interaction effects by weather conditions on the association between crash risk and possible roadway and traffic attributes would be considered. To verify the prevalence of spatio-temporal correlation, difference in the results of goodness-of-fit assessment of the proposed Bayesian approach and random-effect approaches (with and without spatial correlation) would be compared.

The remainder of the paper is organized as follows: Section 2 describes the crash, traffic, roadway, and weather data used for current study. Section 3 provides the details of model formulations, performance assessment, and the Bayesian estimation process. In addition, results of model estimation are presented in Section 4 and the concluding remarks and policy implications are provided in Section 5 respectively.

2. Data

In this study, the integrated data on crash incidence, geometric design, traffic flow and weather condition of Kaiyang Freeway in Guangdong Province of China in 2014 are used. For instances, the crash data are available from the Highway Maintenance and Administration Management Platform maintained by the Guangdong Transportation Group. For the road geometry, information on the horizontal and longitudinal profiles is maintained by the Guangdong Province Communication Planning and Design Institute Company Limited. For the traffic and weather conditions, the traffic count data are extracted from the Guangdong Freeway Network Toll System, and the weather data are collected from the Meteorological Information Management System maintained by the Guangdong Climate Center respectively.

Kaiyang Freeway is of 125 km long four-lane double carriageway highway. It connects three major counties, i.e., Kaiping, Enping, and Yangjiang, in the Guangdong Province. The speed limit is 120 km/h. In this study, the roadway under investigation is stratified into 154 consecutive segments. For every segment, both the horizontal (i.e., curvature) and vertical (i.e., gradient, sag and crest curves, etc.) alignments should be homogeneous. This segmentation approach is consistent to that of the previous studies by the same research team (Wen et al., 2018, 2019; Zeng et al., 2017a). In 2014, there were 692 crashes on Kaiyang Freeway. Each of the 692 crashes is mapped to the corresponding observation unit (by month and segment) using the Geographical Information System (GIS) approach.

In accordance to the standards specified in the Design Specification for Highway Alignment (2006), the values of a number of geometric design and traffic control attributes including pavement type, number of lane, lane width, type and width of median-barrier, type and width of road shoulder and speed limit are fixed for a freeway. The same principle also applies for Kaiyang Freeway. The unique characteristics of each of the 154 roadway segments are the horizontal curvature and vertical gradient. We shed the light from previous studies (Ahmed et al., 2011; Shankar et al., 2004), both the horizontal curvature ('1' denotes a curved segment, while '0' denotes a straight segment respectively) and vertical gradient ('1' denotes a segment of absolute vertical gradient greater than 1%, while '0' otherwise) are stratified into two classes respectively.

For the traffic data, the vehicles are categorized into five classes for the Guangdong Freeway Network Toll System, in accordance to the height, number of axles, number of wheels and wheelbase. Table 1 shows the specifications of each of the five different vehicle types. Taking into account the difference in the maneuver performances between different vehicle types, the overall traffic volume is the weighted sum of that of every vehicle class. Weight of Class 1, 2, 3, 4 and 5 vehicles is 1, 1.5, 2, 3 and 3.5 respectively, in accordance to the specifications set out by the Guangdong Transportation Department. In this study, the total monthly vehicle-kilometers traveled is used as the exposure variable for the proposed crash prediction models. To control for the effect of vehicle composition on the crash risk, proportion of every vehicle class is also incorporated into the proposed model, with which Class 1 vehicle is set as the 'reference'. Also, Class 5 vehicle is not

Table 1
Vehicle classification.

Class	Characteristics				Examples
	Height (m)	Number of axles	Number of wheels	Wheelbase (m)	
1	< 1.3	2	2-4	< 3.2	Passenger car, jeep, pick-up truck
2	≥1.3	2	4	≥3.2	Minibus, minivan, light truck
3	≥1.3	2	6	≥3.2	Medium bus, large ordinary bus, medium truck
4	≥1.3	3	6-10	≥3.2	Large luxury bus, large truck, large trailer, 20-feet container truck
5	≥1.3	> 3	> 10	≥3.2	Heavy truck, heavy trailer, 40-feet container truck

considered since there is significant correlation between Class 4 and Class 5 vehicles (Zeng et al., 2017a).

In this study, the meteorological data from three county-level weather stations: (i) Kaiping Station; (ii) Enping Station; and (iii) Yangjiang Station, are extracted from the Meteorological Information Management System. The freeway segments under investigation are assigned to one of the three weather stations in accordance to the distance between the segment and weather stations. Meteorological variables considered are wind speed, precipitation, and visibility. Information on the daily average wind speed, visibility, and precipitation are available. These data are aggregated into monthly average, and incorporated into the proposed crash prediction models. Table 2 shows the descriptive statistics of the variables considered.

3. Methodology

In this study, the fit performance of three candidate models: Random-effect model; Spatial model; and Spatio-temporal model is assessed. The Bayesian estimation is conducted using WinBUGS.

3.1. Model specification

3.1.1. Random effect model

Random effect model (also known as "hierarchical Poisson model") is a common method for Bayesian hierarchical modeling. The observed crash count $Y_{i,t}$ on freeway segment i $(i = 1,2, \ldots,154)$ in month t $(t = 1,2, \ldots,12)$ is assumed to follow a Poisson distribution (Lord et al., 2005):

$$P(Y_{i,t} = k) = \frac{(\lambda_{i,t})^k}{k!}\exp(-\lambda_{i,t}),\ k = 0,1,\ 2,\ \ldots \tag{1}$$

in which $\lambda_{i,t}$ is the expected crash count and is formulated as:

$$\ln\lambda_{i,t} = \ln e_{i,t} + \boldsymbol{\beta}\mathbf{X}_{i,t} + \theta_i \tag{2}$$

Table 2
Descriptive statistics of the variables used for the analysis.

Variables	Description	Mean	S.D.	Min.	Max.
RESPONSE VARIABLE					
Crash	Monthly total crash count	0.38	0.68	0	5
EXPOUSRE VARIABLE					
MVKT	Monthly vehicle kilometer travelled (10^3 km·pcu[a])	1349	524	272	4593
TRAFFIC COMPOSITION					
Veh_3	Proportion of Class 3 vehicle (%)	21.9	2.15	15.7	25.5
Veh_4	Proportion of Class 4 vehicle (%)	6.59	0.77	4.57	7.80
ROADWAY CHARACTERISTICS					
Curve	1 = Curved road, 0 = otherwise	0.75	0.44	0	1
Slope	1 = Absolute gradient > 1%, 0 = otherwise	0.25	0.44	0	1
WEATHER CONDITION					
Wind speed	Average monthly wind speed (m/s)	2.49	0.88	1.53	4.57
Precipitation	Monthly average daily precipitation (mm)	4.48	3.67	0.003	13.0
Visibility	Average monthly visibility (km)	20.5	11.4	7.71	39.8

[a] pcu: passenger car unit.

In Eq. (2), $e_{i,t}$ is the crash exposure on freeway segment i during period t and is defined as a power of MVKT (Yu et al., 2013). $\boldsymbol{\beta}$ is a vector of parameters corresponding to that of explanatory variables (including a constant element), $\mathbf{X}_i$. θ_i is used to capture the site-specific random effects shared by the observations (within and between months) and is able to handle the over-dispersion issue (Ahmed et al., 2011). It is assumed to follow a normal distribution with zero mean and $\sigma_h(>0)$ variance:

$$\theta_i \tilde{}\ Normal(0,\ \sigma_h) \tag{3}$$

3.1.2. Spatial model

The observation units are consecutive segments on the same freeway, the segments in close proximity could share similar (unobserved) factors, i.e. lighting condition, vehicle platoon, pavement condition, and roadside features, etc., that may affect the crash incidence. This could result in bias in parameter estimation attributed to spatial correlation (Zeng and Huang, 2014). To account for the effect of spatial correlation, as proposed by Besag (1974), a residual term φ_i with conditional autoregressive (CAR) prior is added into the link function:

$$\ln\lambda_{i,t} = \ln e_{i,t} + \boldsymbol{\beta}\mathbf{X}_{i,t} + \theta_i + \varphi_i \tag{4}$$

$$\varphi_i \tilde{}\ Normal\left(\frac{\sum_{i\neq j}\omega_{i,j}\varphi_j}{\sum_{i\neq j}\omega_{i,j}},\ \frac{\sigma_s}{\sum_{i\neq j}\omega_{i,j}}\right) \tag{5}$$

where $\omega_{i,j}$ is the adjacency weight between freeway segments i and j. The most prevalent first order neighbor structure is used in the current research to define the adjacency weights (Aguero-Valverde and Jovanis, 2008): $\omega_{i,j} = 1$ if segments i and j are mutually connected, and $\omega_{i,j} = 0$ otherwise. $\sigma_s(>0)$ is the variance parameter of spatial correlation.

The posterior proportion of extra-Poisson variation accounted for by the spatial correlation is of interest and is calculated as:

$$\alpha_s = \frac{sd(\boldsymbol{\varphi})}{sd(\boldsymbol{\varphi}) + sd(\boldsymbol{\theta})},\ \boldsymbol{\varphi} = (\varphi_1,\ \varphi_2, \ldots, \varphi_{154}),\ \boldsymbol{\theta} = (\theta_1,\ \theta_2,\ \ldots, \theta_{154}) \tag{6}$$

3.1.3. Spatio-temporal model

Previous study by the same research team (Zeng et al., 2017a) indicates that a residual term $\delta_{i,t}$ with lag-1 dependence is an appropriate method for accommodating the temporal correlation among crash counts in consecutive time periods. Lag-1 implies that the temporal effect on a specific freeway section in current period is affected by its counterpart in the previous one (month in current study). Based on the stationarity assumption (Congdon, 2014), it is assumed to be normally distributed as followings:

$$\delta_{i,1} \tilde{}\ Normal\left(0,\ \frac{\sigma_l}{1-\gamma^2}\right) \tag{7}$$

$$\delta_{i,t} \tilde{}\ Normal(\gamma\delta_{i,1},\ \sigma_l) \tag{8}$$

in which γ is the autocorrelation coefficient and σ_l is the variance parameter of temporal correlation.

Moreover, the interaction between spatial and temporal effects is taken into consideration because there may be variations in the spatial or temporal correlation across months or freeway segments, respectively (Dong et al., 2016). The spatio-temporal interaction is formulated as the product of a time trend t and a spatial term ϕ_i with CAR prior:

$$\phi_i \tilde{} Normal\left(\frac{\sum_{i\neq j}\omega_{i,j}\phi_j}{\sum_{i\neq j}\omega_{i,j}}, \frac{\sigma_a}{\sum_{i\neq j}\omega_{i,j}}\right) \quad (9)$$

where $\sigma_a(>0)$ is the variance parameter of the spatial component of the spatio-temporal interaction.

Hence, the spatio-temporal model's link function is expressed as:

$$\ln\lambda_{i,t} = \ln e_{i,t} + \boldsymbol{\beta}\mathbf{X}_{i,t} + \theta_i + \varphi_i + \delta_{i,t} + t\phi_i \quad (10)$$

In the spatio-temporal model, we redefine the proportion accounted for by the spatial correlation in Eq. (11), and also calculate the proportions explained by the temporal correlation and spatio-temporal interaction in Eqs. (12) and (13) respectively:

$$\alpha_s = \frac{sd(\boldsymbol{\varphi})}{sd(\boldsymbol{\varphi}) + sd(\boldsymbol{\delta}) + sd(\mathbf{I}) + sd(\boldsymbol{\theta})} \quad (11)$$

$$\alpha_t = \frac{sd(\boldsymbol{\delta})}{sd(\boldsymbol{\varphi}) + sd(\boldsymbol{\delta}) + sd(\mathbf{I}) + sd(\boldsymbol{\theta})} \quad (12)$$

$$\alpha_I = \frac{sd(\mathbf{I})}{sd(\boldsymbol{\varphi}) + sd(\boldsymbol{\delta}) + sd(\mathbf{I}) + sd(\boldsymbol{\theta})} \quad (13)$$

where $\boldsymbol{\delta} = (\delta_{1,1}, \delta_{1,2}, ..., \delta_{154,12})$, $\mathbf{I} = (\phi_1, 2\phi_1, ..., 12\phi_{154})$.

3.2. Goodness-of-fit measures

Three commonly used criteria, the deviance information criterion (DIC), mean absolute deviance (MAD), and mean squared prediction error (MSPE), are used in current study to measure the goodness-of-fit of the proposed Bayesian models. According to the definition proposed by Spiegelhalter et al. (2002), DIC is calculated as:

$$DIC = \bar{D} + pD \quad (14)$$

in which $\bar{D}$ is the posterior mean deviance for assessing the model fit, and pD is the effective number of parameters in the model and is used for measuring model complexity. Generally, a model with a lower DIC value is preferred. DIC differences between 5 and 10 are deemed substantial, while differences more than 10 indicate the significant outperformance of the model with a lower DIC (Spiegelhalter et al., 2005).

MAD and MSPE are defined in the following two equations respectively (Zeng and Huang, 2014):

$$MAD = \frac{1}{12 \times 154}\sum_{i=1}^{154}\sum_{t=1}^{12}|Y_{i,t} - \lambda_{i,t}| \quad (15)$$

$$MSPE = \frac{1}{12 \times 154}\sum_{i=1}^{154}\sum_{t=1}^{12}(Y_{i,t} - \lambda_{i,t})^2 \quad (16)$$

3.3. Model estimation

All the three models are coded and diagnosed using WinBUGS (Lunn et al., 2000), which provides a friendly environment for Bayesian inference using Markov chain Monte Carlo (MCMC) simulation approach with Gibbs sampling algorithm. Without sufficient prior knowledge, non-informative priors are specified for the parameters and hyperparameters. Specifically, we use a diffused normal distribution $\mathrm{Normal}(0, 10^4)$ as the priors of the regression coefficients (i.e., the elements of $\boldsymbol{\beta}$) and the autocorrelation coefficient γ (Zeng et al., 2018). The CAR priors are specified by the *car.normal* function available in WinBUGS. A diffused gamma distribution, $\mathrm{Gamma}(0.001, 0.001)$, is used as the priors of the precision parameters (i.e., the reciprocal of the variance parameters, $1/\sigma_h$, $1/\sigma_s$, $1/\sigma_l$, and $1/\sigma_a$) for site-specific, spatial, temporal, and spatio-temporal interaction effects (Dong et al., 2016).

Table 3
Results of model estimation.

	Random effect model	Spatial model	Spatio-temporal model
Constant	−8.60 (−11.0, −6.08)[a]	−6.91 (−9.31, −4.10)	−7.00 (−9.74, −4.21)
MVKT	1.13 (0.86, 1.44)	1.03 (0.73, 1.28)	1.05 (0.73, 1.34)
Veh_4	−0.10 (−0.21, 0.01)[b]	−0.12 (−0.23, −0.01)	−0.15 (−0.26, −0.03)
Slope	−1.48 (−2.23, −0.73)	−1.51 (−2.32, −0.71)	−1.47 (−2.31, −0.65)
*Wind speed*Slope*	0.52 (0.27, 0.77)	0.51 (0.24, 0.77)	0.49 (0.21, 0.78)
*Precipitation*Curve*	0.10 (0.03, 0.18)	0.10 (0.02, 0.18)	0.09 (0.02, 0.17)
*Precipitation*Slope*	−0.08 (−0.15, −0.01)	−0.08 (−0.15, −0.01)	−0.08 (−0.16, −0.01)
*Visibility*Slope*	0.04 (0.01, 0.06)	0.04 (0.01, 0.06)	0.04 (0.01, 0.06)
$sd(\boldsymbol{\theta})$	0.22 (0.06, 0.36)	0.17 (0.03, 0.33)	0.11 (0.02, 0.28)
$sd(\boldsymbol{\varphi})$	—	0.29 (0.10, 0.48)	0.31 (0.08, 0.53)
$sd(\boldsymbol{\delta})$	—	—	0.33 (0.05, 0.59)
$sd(\mathbf{I})$	—	—	0.29 (0.18, 0.41)
α_s	—	0.63 (0.29, 0.91)	0.30 (0.10, 0.48)
α_t	—	—	0.31 (0.09, 0.50)
α_I	—	—	0.28 (0.17, 0.42)
$\bar{D}$	2792	2780	2677
pD	42	45	137
DIC	2834	2825	2814
MAD	0.48	0.47	0.45
MSPE	0.41	0.41	0.38

[a] Estimated mean (95% Bayesian credible interval) for the parameter.
[b] Insignificant at the 95% credibility level.

For each model, we run a chain of 100,000 MCMC simulation iterations, with the first 50,000 iterations acting as a burn-in. To determine the simulation convergence, we monitor the ratios between the Monte Carlo simulation errors and the respective estimates' standard deviations to ensure that they are less than 0.05 (Spiegelhalter et al., 2005). The results of model estimation and comparison are shown in Table 3, where only the variables statistically significant (at the 95% credibility level) at least in one model are included. In this study, we intend to examine the interaction effects by weather condition on the association between crash frequency and possible roadway characteristics. For instances, six interaction terms: (i) *Curve* × *Wind speed*; (ii) *Curve* × *Precipitation*; (iii) *Curve* × *Visibility*; (iv) *Slope* × *Wind speed*; (v) *Slope* × *Precipitation*, and (vi) *Slope* × *Visibility,* are considered.

4. Results

4.1. Model comparison

We first compare the spatial model and random effect model. The DIC value of the spatial model (2825) is lower than that of the random effect model (2834). This indicates that the spatial model is superior. Such a result is in line with that of the previous studies (Zeng and Huang, 2014; Zeng et al., 2019). Indeed, taking into account the effect of spatial correlation in crash data using CAR priors can mitigate the model misspecification and improve model fit. Significant spatial effect is also confirmed by the Bayesian estimate of the spatial term's standard deviation, $sd(\boldsymbol{\varphi}) = 0.29$, which accounts for 63% of the variation. The spatial correlation could be attributed to the effects of unobserved factors that are spatially clustered, such as terrain feature and lighting condition (Wen et al., 2018, 2019). After accounting for spatial correlation, the magnitude of random effect θ_i reduces from 0.22 to 0.17. The result is reasonable, as the previous studies (Barua et al., 2016; Zeng and Huang, 2014) have pointed out that the unstructured random effects may be derived from several potential sources including spatial correlation.

Among the three candidate models, the spatio-temporal model fits with the observed outcomes the best, because it yields the lowest DIC

(2814), MAD (0.45), and MSPE (0.38). Such finding is anticipated, because the temporal term specified in lag-1 structure can proxy the lag effect of the traffic and weather conditions in the preceding time period (Cheng et al., 2018a), while the spatio-temporal term can accommodate the temporal variation of spatial effect or spatial variation of temporal effect (Dong et al., 2016). The estimated standard deviations, $sd(\boldsymbol{\delta})$ and $sd(\mathbf{I})$, reveal that there are also remarkable temporal correlation and spatio-temporal interaction in the crash data, which explain 31% and 28% of the variation respectively, as suggested by the estimates of α_t and α_I. In addition, the magnitude of spatial correlation is comparable to the counterpart in the spatial model (see Table 3), but the proportion of variation attributed to it reduces to 30% only. The minimal unstructured random effect (0.11) of this model implies that temporal correlation and spatio-temporal interaction could be influential.

Results indicate that coefficient estimates are generally consistent among the spatial model and spatio-temporal model (same sign and comparable posterior means), but not that of the random effect model. For example, the coefficient of *Veh_4* is found to be significant (at the 95% credibility level) only in the spatial model and spatio-temporal model. Many other studies (Dinu and Veeraragavan, 2011; Wen et al., 2018; Zeng et al., 2017a) have reported that traffic composition is an important contributory factor to crash frequency. Such a finding also indicates that the models that account for the spatial and/or temporal correlation and their interaction are superior.

4.2. Parameter estimates

Purpose of this study is to examine the effects of road geometry, traffic characteristics, weather conditions and their interactions on the freeway crash incidence. The above-mentioned results indicate that the spatio-temporal model is superior.

As shown in Table 3, increase in vertical gradient (greater than 1%) is correlated to the reduction in crash frequency (Coefficient = -1.47). This apparently contradicts that of previous studies (Ahmed et al., 2011; Wen et al., 2018; Zeng et al., 2017a). However, this could be attributed to the difference in the reporting rates across different areas (i.e. urban versus rural) and road types (with different standards governing the horizonal and vertical alignments), attributed to the speculated attitude of the driver for the crash responsibility at the challenging terrain (Tay et al., 2009).

Results indicate that the interaction between *slope* and *wind speed* on crash frequency is statistically significant. This indicates that the negative impact of wind speed on the crash risk could magnify with the vertical gradient. This finding is consistent to that of previous studies (Hou et al., 2018b; Young and Liesman, 2007). It could be attributed to the increase in the risk of overturning and instability under strong wind (and wind gusts) on the slopes.

In addition, the interaction between *curve* and *precipitation*, is found to have a significantly positive impact on crash frequency. It could be attributed to the reduction in skid resistance when the precipitation is accumulated on the pavement surface (Zeng et al., 2016). This could in turn increase the centripetal force required for the vehicle to negotiate a curve, and therefore increase the risk of overturning and running-off.

On the other hand, the interactions between *slope* and *precipitation*, and between *slope* and *visibility*, also significantly impact the crash incidence. In particular, increase in precipitation could mitigate the negative impact of vertical gradient on crash risk, while the increase in visibility could magnify the negative impact of vertical gradient. These could be attributed to the risk compensation of the driver (Mannering and Bhat, 2014). In particular, drivers could be more cautious when driving under the adverse weather conditions (i.e., heavy rain and poor visibility). Therefore, it reduces the crash risk in difficult terrain (e.g., steep slope) (Zeng et al., 2017b).

For the traffic condition, the exposure *MVKT* is found positively correlated to the crash incidence at the 95% credibility level. The result is intuitive and in line with that of previous studies (Wen et al., 2018; Yu et al., 2013). Nevertheless, results of Bayesian estimate indicate that the relationship between crash frequency and vehicle kilometers traveled was linear. For the traffic composition, increase in the proportion of heavy vehicle (Class 4) is correlated to the reduction in the crash risk. Similar results can be found in our previous study (Zeng et al., 2017a). This could be attributed to: (1) the better driving skills of the professional drivers (Class 4) than that of the passenger car (Class 1 as the reference); (2) the professional drivers tend to be more familiar with the freeway because they have been traveling on the freeway routinely; (3) working and driving time regulations imposed on the professional drivers (e.g., the heavy vehicle drivers are not allowed to drive for more than four hours in China in the night time); (4) more trucks in the traffic may increase the uniformity of the platoons and thus reduce the probability of traffic conflicts.

5. Conclusions and future research

This study investigates the effects of roadway characteristics (i.e., horizontal and vertical alignments), traffic and weather conditions (i.e., wind speed, precipitation, and visibility), and their interactions on the freeway crash incidence, using the crash data on Kaiyang Freeway in China in 2014. A Bayesian spatio-temporal model is proposed to model the association, taking into account the effects of site-specific random effect, spatial correlation across neighboring freeway segments, temporal correlation across consecutive time periods, and the spatio-temporal interaction.

Results of the Bayesian spatio-temporal model indicate that increase in the wind speed, reduction in the precipitation, and increase in the visibility could magnify the negative safety impact of vertical gradient. In addition, increase in the precipitation could magnify the negative impact of horizontal curves. Increase in the proportion of heavy vehicle could mitigate the negative impact of vertical gradient. The above could be attributed to the risk compensation behavior of the driver, and the driving skills of professional drivers. Such findings are indicative of the development of real-time traffic management and control measures that could enhance the safety performance of the freeway. For example, variable message signs could be installed at the curved road sections and difficult terrains, whereas the warning could be disseminated under the adverse weather conditions, such as heavy rainfall and low visibility.

From the perspective of methodological advances, the proposed Bayesian model could take into account the effects of spatial correlation, temporal correlation and spatio-temporal interaction on the parameter estimation, as indicated by the low DIC, MAD and MSPE values, and the considerable variations to be explained by the correlation term, given that the time-series crash, traffic and weather data are extracted from the consecutive freeway segments on the same freeway. The unobserved traffic and environmental factors could be correlated with the association.

Yet, the effects of only two roadway factors on the freeway crash risk are explored in current study. It would be worth exploring the effects of other geometric design characteristics, e.g., number of lanes, lane width, and shoulder width, and their interactions with the weather conditions, on the crash risk, when the comprehensive traffic and crash data of a bigger freeway network are available in the extended study. Methodological-wise, it is a common approach to examine the interaction effects on the association by adding the corresponding interaction terms into the link function of regression. Yet, it is worth exploring the viability of other alternate approaches, e.g., machine learning techniques (Zeng et al., 2016), to reveal the interaction effects when more comprehensive traffic, driver and weather data are available in future study. Additionally, a random-parameter model can be set out to capture the heterogeneous effects of the observed factors on the association (Mannering et al., 2016).

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.

Acknowledgement

This study was jointly supported by: 1) the Natural Science Foundation of China under Grant Nos. 51578247and71801095; 2) the Natural Science Foundation of Guangdong Province under Grant No.2017A030310161; 3) the Research Grants Council of Hong Kong (Project No. 25203717); and 4) the International Science & Technology Cooperation Program of China under Grant No. SQ2017YFGH002253.

References

Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accid. Anal. Prev. 38 (3), 618–625.
Aguero-Valverde, J., Jovanis, P.P., 2008. Analysis of road crash frequency with spatial models. Transp. Res. Rec. 2061, 55–63.
Ahmed, M., Franke, R., Ksaibati, K., Shinstine, D.S., 2018. Effects of truck traffic on crash injury severity on rural highways in Wyoming using Bayesian binary logit models. Accid. Anal. Prev. 117, 106–113.
Ahmed, M., Huang, H., Abdel-Aty, M., Guevara, B., 2011. Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway. Accid. Anal. Prev. 43 (4), 1581–1589.
Barua, S., El-Basyouny, K., Islam, M.T., 2016. Multivariate random parameters collision count data models with spatial heterogeneity. Anal. Methods Accid. Res. 9, 1–15.
Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. J. R. Stat. Soc. Ser. B 36 (2), 192–225.
Cheng, W., Gill, G.S., Zhang, Y., Cao, Z., 2018a. Bayesian spatiotemporal crash frequency models with mixture components for space-time interactions. Accid. Anal. Prev. 112, 84–93.
Cheng, W., Gill, G.S., Ensch, J.L., Kwong, J., Jia, X., 2018b. Multimodal crash frequency modeling: multivariate space-time models with alternate spatiotemporal interactions. Accid. Anal. Prev. 113, 159–170.
Congdon, P., 2014. Applied Bayesian Modelling. John Wiley & Sons.
De Luca, M., Dell'Acqua, G., 2012. Freeway safety management: case studies in Italy. Transport 27 (3), 320–326.
Design Specification for Highway Alignment, 2006. Ministry of Transport of the People's Republic of China (JTG D20-2006). .
Dinu, R.R., Veeraragavan, A., 2011. Random parameter models for accident prediction on two-lane undivided highways in India. J. Safety Res. 42 (1), 39–42.
Dong, N., Huang, H., Lee, J., Gao, M., Abdel-Aty, M., 2016. Macroscopic hotspots identification: a Bayesian spatio-temporal interaction approach. Accid. Anal. Prev. 92, 256–264.
Gaweesh, S.M., Ahmed, M.M., Piccorelli, A.V., 2019. Developing crash prediction models using parametric and nonparametric approaches for rural mountainous freeways: a case study on Wyoming Interstate 80. Accid. Anal. Prev. 123, 176–189.
Hou, Q., Meng, X., Leng, J., Yu, L., 2018a. Application of a random effects negative binomial model to examine crash frequency for freeways in China. Phys. A Stat. Mech. Its Appl. 509, 937–944.
Hou, Q., Tarko, A.P., Meng, X., 2018b. Investigating factors of crash frequency with random effects and random parameters models: new insights from Chinese freeway study. Accid. Anal. Prev. 120, 1–12.
Hou, Q., Meng, X., Huo, X., Cheng, Y., Leng, J., 2019. Effects of freeway climbing lane on crash frequency: application of propensity scores and potential outcomes. Phys. A Stat. Mech. Its Appl. 517, 246–256.
Huang, H., Song, B., Xu, P., Zeng, Q., Lee, J., Abdel-Aty, M., 2016. Macro and micro models for zonal crash prediction with application in hot zones identification. J. Transp. Geogr. 54, 248–256.
Jonkers, E., Wilmink, I.R., Stoelhorst, H., Schreuder, M., Polderdijk, S., 2011. Results of field trials with dynamic speed limits in the Netherlands: improving throughput and safety on the A12 freeway. October. 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC). pp. 2168–2173.
Jung, S., Jang, K., Yoon, Y., Kang, S., 2014. Contributing factors to vehicle to vehicle crash frequency and severity under rainfall. J. Safety Res. 50, 1–10.
Labi, S., 2011. Efficacies of roadway safety improvements across functional subclasses of rural two-lane highways. J. Safety Res. 42 (4), 231–239.
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.
Lord, D., Washington, S.P., Ivan, J.N., 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accid. Anal. Prev. 37 (1), 35–46.
Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS-A Bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. 10 (4), 325–337.
Lyu, N., Cao, Y., Wu, C., Xu, J., Xie, L., 2018. The effect of gender, occupation and experience on behavior while driving on a freeway deceleration lane based on field operational test data. Accid. Anal. Prev. 121, 82–93.
Ma, X., Chen, S., Chen, F., 2017a. Multivariate space-time modeling of crash frequencies by injury severity levels. Anal. Methods Accid. Res. 15, 29–40.
Ma, Z., Zhang, H., Steven, I., Chien, J., Wang, J., Dong, C., 2017b. Predicting expressway crash frequency using a random effect negative binomial model: a case study in China. Accid. Anal. Prev. 98, 214–222.
Malyshkina, N.V., Mannering, F.L., Tarko, A.P., 2009. Markov switching negative binomial models: an application to vehicle accident frequencies. Accid. Anal. Prev. 41 (2), 217–226.
Mannering, F.L., Bhat, C.R., 2014. Analytic methods in accident research: methodological frontier and future directions. Anal. Methods Accid. Res. 1, 1–22.
Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.
Ponnaluri, R.V., 2016. The odds of wrong-way crashes and resulting fatalities: a comprehensive analysis. Accid. Anal. Prev. 88, 105–116.
Sarhan, M., Hassan, Y., Abd El Halim, A.O., 2008. Safety performance of freeway sections and relation to length of speed-change lanes. Can. J. Civ. Eng. 35 (5), 531–541.
Shankar, V., Chayanan, S., Sittikariya, S., Shyu, M.B., Juvva, N., Milton, J., 2004. Marginal impacts of design, traffic, weather, and related interactions on roadside crashes. Transp. Res. Rec. 1897, 156–163.
Shew, C., Pande, A., Nuworsoo, C., 2013. Transferability and robustness of real-time freeway crash risk assessment. J. Safety Res. 46, 83–90.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. Series B Stat. Methodol. 64 (4), 583–639.
Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2005. WinBUGS User Manual. MRC Biostatistics Unit, Cambridge.
Tay, R., Barua, U., Kattan, L., 2009. Factors contributing to hit-and-run in fatal crashes. Accid. Anal. Prev. 41 (2), 227–233.
Wen, H., Sun, J., Zeng, Q., Zhang, X., Yuan, Q., 2018. The effects of traffic composition on freeway crash frequency by injury severity: A Bayesian multivariate spatial modeling approach. J. Adv. Transp. https://doi.org/10.1155/2018/6964828. Article ID: 6964828,.
Wen, H., Zhang, X., Zeng, Q., Lee, J., Yuan, Q., 2019. Investigating spatial autocorrelation and spillover effects in freeway crash-frequency data. Int. J. Environ. Res. Public Health 16 (2), 219.
Wood, J.S., Donnell, E.T., Fariss, C.J., 2016. A method to account for and estimate underreporting in crash frequency research. Accid. Anal. Prev. 95, 57–66.
Wu, Y., Abdel-Aty, M., Lee, J., 2018. Crash risk analysis during fog conditions using real-time traffic data. Accid. Anal. Prev. 114, 4–11.
Ye, X., Pendyala, R.M., Shankar, V., Konduri, K.C., 2013. A simultaneous equations model of crash frequency by severity level for freeway sections. Accid. Anal. Prev. 57, 140–149.
Young, R.K., Liesman, J., 2007. Estimating the relationship between measured wind speed and overturning truck crashes using a binary logit model. Accid. Anal. Prev. 39 (3), 574–580.
Yu, R., Abdel-Aty, M., 2013. Multi-level Bayesian analyses for single-and multi-vehicle freeway crashes. Accid. Anal. Prev. 58, 97–105.
Yu, R., Abdel-Aty, M., Ahmed, M., 2013. Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors. Accid. Anal. Prev. 50, 371–376.
Zeng, Q., Huang, H., 2014. Bayesian spatial joint modeling of traffic crashes on an urban road network. Accid. Anal. Prev. 67, 105–112.
Zeng, Q., Huang, H., Pei, X., Wong, S.C., 2016. Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks. Anal. Methods Accid. Res. 10, 12–25.
Zeng, Q., Sun, J., Wen, H., 2017a. Bayesian hierarchical modeling monthly crash counts on freeway segments with temporal correlation. J. Adv. Transp. https://doi.org/10.1155/2017/5391054. Article ID: 5391054,.
Zeng, Q., Wen, H., Huang, H., Abdel-Aty, M., 2017b. A Bayesian spatial random parameters Tobit model for analyzing crash rates on roadway segments. Accid. Anal. Prev. 100, 37–43.
Zeng, Q., Wen, H., Huang, H., Pei, X., Wong, S.C., 2018. Incorporating temporal correlation into a multivariate random parameters Tobit model for modeling crash rate by injury severity. Transp. A Transp. Sci. 14 (3), 177–191.
Zeng, Q., Gu, W., Zhang, X., Wen, H., Lee, J., Hao, W., 2019. Analyzing freeway crash severity using a Bayesian generalized spatial ordered logit model with conditional autoregressive priors. Accid. Anal. Prev. 127, 87–95.
Zhai, X., Huang, H., Sze, N.N., Song, Z., Hon, K.K., 2019. Diagnostic analysis of the effects of weather condition on pedestrian crash severity. Accid. Anal. Prev. 122, 318–324.
Zou, Y., Zhang, Y., Lord, D., 2013. Application of finite mixture of negative binomial regression models with varying weight parameters for vehicle crash data analysis. Accid. Anal. Prev. 50, 1042–1051.