Analytic Methods in Accident Research 28 (2020) 100137

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

Modeling pedestrian-injury severities in pedestrian-vehicle crashes considering spatiotemporal patterns: Insights from different hierarchical Bayesian random-effects models

Li Song [a], Yang Li [a], Wei (David) Fan [a,*], Peijie Wu [b]

[a] USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE), Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, EPIC Building, Room 3366, 9201 University City Boulevard, Charlotte, NC 28223-0001, United States
[b] School of Transportation Science and Engineering, Harbin Institute of Technology, No. 73, Huanghe Road, Nangang District, Harbin, China

ARTICLE INFO

Article history:
Received 16 August 2020
Received in revised form 13 October 2020
Accepted 13 October 2020
Available online 20 October 2020

Keywords:
Pedestrian
Injury severity
Spatiotemporal analysis
Hierarchical Bayesian model
Random effects only

ABSTRACT

To systematically account for the spatiotemporal features and unobserved heterogeneity within pedestrian-vehicle crashes, this paper employs the spatiotemporal analysis and hierarchical Bayesian random-effects models to explore the factors contributing to pedestrian-injury severities of pedestrian-vehicle crashes involving single vehicle in North Carolina from 2007 to 2018. Ten spatiotemporal patterns of the crashes are identified by applying an improved spatiotemporal analysis. Significant temporal instability and the spatiotemporal instability of the factors to the pedestrian-injury crashes are identified by the likelihood ratio tests. A hierarchical Bayesian random intercept logit model with random-effects across the spatiotemporal groups is firstly employed for the whole dataset. The comparison between different hierarchical models indicates that addressing random-effects across observations and increasing the number of random parameters could both improve the model performance. Then a hierarchical Bayesian random-effects-only logit model, which allows all parameters to be randomly distributed across observations, is developed to further investigate the unobserved heterogeneity in spatiotemporal segmented datasets. The significant improvements in terms of model fit and the hit accuracy underscore the superiority of the random-effects-only model. The marginal effects of the human, vehicle, crash, locality, roadway, environment, time, and traffic control factors for each spatiotemporal dataset also provide insights into possible inherent reasons for the spatiotemporal instability/tendency of the crash and correlated factors. Meanwhile, specific countermeasures are given to locations especially in which the spatially aggregated patterns of the crashes have new, consecutive, and intensifying temporal tendencies. This study provides a framework for engineers and researchers to identify spatiotemporal patterns of the crashes and explore the factors affecting pedestrian-injury severities especially in those existing crash-prone areas.

* Corresponding author.
E-mail addresses: lsong1@uncc.edu (L. Song), yli107@uncc.edu (Y. Li), wfan7@uncc.edu (Wei (David) Fan), peijiewu@foxmail.com (P. Wu).

https://doi.org/10.1016/j.amar.2020.100137
2213-6657/

1. Introduction

Pedestrians are more likely to suffer severe injuries than other road users in traffic crashes due to the lack of protection, mass, and speed (Li et al., 2017). Beck et al. (2007) also estimated that the fatality rate per trip for pedestrians is 1.5 times larger than passenger vehicle occupants. From 2008 to 2017, the percentage of pedestrian deaths in total traffic fatalities in the U.S. has increased from 12% to 16% (NHTSA, 2019). Hence, pedestrian-injury severities have been the emphasis of several research efforts over the years. The ongoing methodological frontiers of the crash injury severities research are mainly focused on accounting for the unobserved heterogeneity and the spatial/temporal patterns of the factors affecting crash severities (Behnood and Mannering, 2016; Mannering and Bhat, 2014). Mannering (2018) pointed out several human behavioral reasons for the temporal instability of the crash factors and the ignorance of the temporal nature would result in inaccurate model estimations and ineffective development of safety countermeasures.

Except for the temporal features of pedestrian-vehicle crashes, their spatial characteristics should also be considered simultaneously (Liu et al., 2019). The pedestrian-vehicle crashes in some locations show significant aggregated/dispersed patterns. Meanwhile, these spatial patterns may also change over time. For instance, some locations have aggregated crashes only in the past but not in recent years. However, the spatiotemporal patterns of the pedestrian-vehicle crashes have not been adequately studied in the past. According to Behnood and Mannering (2016), the existence of the temporal trends in the pedestrian-vehicle crash might impact the model results, and neglecting these temporal patterns could result in incorrect conclusions (Mannering, 2018). A thorough study of the spatiotemporal patterns of the pedestrian crashes might provide insights into possible inherent reasons for the spatiotemporal instability of the crashes and factor effects.

As the Bayesian reference is especially suitable for estimating random-effects in complex hierarchical models, some previous studies employed hierarchical Bayesian random-effects models to account for unobserved heterogeneity (Chen et al., 2016; Huang et al., 2008; Mitra and Washington, 2007). Additionally, since the comparisons between different hierarchical Bayesian models have not been thoroughly addressed, this paper also explores the differences between hierarchical Bayesian random-effects models. First, this paper compares the models that account for heterogeneity across the observations and groups. Then, this paper examines the models that address the random-effects in intercept terms, in individual characteristics variables, and in all variables, respectively. For example, a random-intercept model is proposed for the whole dataset to account for the heterogeneity across the spatiotemporal groups. Also, a random-effects-only model is developed for spatiotemporal segmented datasets to further account for unobserved heterogeneity across observations. Since all parameters are treated as random parameters across observations, the random-effects-only model is supposed to significantly improve the model fit and prediction accuracy.

By using pedestrian-vehicle crash data in North Carolina, this paper explores the factors affecting pedestrian-injury severities considering the unobserved heterogeneity and spatiotemporal patterns of the crash data. An improved spatiotemporal analysis method is utilized to identify spatiotemporal patterns of the crash. Temporal and spatiotemporal instability of the factors are tested by the likelihood ratio tests. Different hierarchical Bayesian random-effects models are employed for the whole dataset and spatiotemporal segmented datasets. Comparisons between models, marginal effects results, and specific countermeasures for spatially aggregated crashes with different temporal patterns are further discussed.

2. Literature review

2.1. Spatiotemporal patterns of the crashes

To identify locations with spatially aggregated/high-frequency crashes, previous studies usually employed point pattern analysis such as the kernel density estimation (KDE) (Ouni and Belloumi, 2018) and the Getis-Ord Gi* index (Songchitruksa and Zeng, 2010). Since KDE is not feasible for the statistical test and the choice of bandwidth would significantly influence the density pattern results (Plug et al., 2011), the Getis-Ord Gi* index, which is a statistical test for aggregated/dispersed clusters, was proposed. Li et al. (2020) compared the KDE with the Getis-Ord Gi and the results indicated that the KDE would identify more high traffic violation points than the Getis-Ord Gi method. However, they also mentioned that it is hard to directly compare the performance of these two methods as the basic principle of these two methods are different. Ulak et al. (2017) employed the Getis-Ord Gi* index to detect crash-prone locations with the neighborhood distance that was calculated by the Global Moran's I test. Results indicated that the distance to hospitals is one of the major reasons for severe injuries on several roadway segments. Meanwhile, to provide a reference for calculating the neighborhood distance and the distance interval, the global Moran's I and the average nearest neighbor (ANN) are further utilized according to (Blazquez et al., 2018; Yalcin and Duzgun, 2015).

For investigations on temporal features of the crashes, Behnood and Mannering (2016) identified a significant temporal instability of the pedestrian-injury severity factors during three economic periods (i.e., pre-recession, recession, and post-recession). It is noted that the spatial patterns of the crashes also have several non-linear temporal features. Thus, to account for non-linear temporal features, the Mann-Kendall trend test, which is a statistical non-parametric rank analysis method, was proposed (Kendall and Gibbons, 1990; Mann, 1945). Gudes et al. (2017) evaluated temporal patterns of the locations with aggregated/dispersed heavy-vehicle crashes by applying the Mann-Kendall trend test. Results indicated inconsistent temporal patterns in aggregated crash locations over time. The Mann-Kendall trend test could identify the temporal patterns

including not only those intensifying and diminishing patterns, but also new emerging, consecutive, persistent, sporadic, oscillating, and historical patterns. And these non-linear patterns are more representative of the realistic temporal patterns of the crashes.

2.2. Pedestrian-injury severity studies considering unobserved heterogeneity

As summarized in Table 1, ordered/unordered discrete outcome models with a logit/probit link function have been widely employed in pedestrian-injury severity studies because of their excellent performance in parameter estimation and outcome inference (Mannering and Bhat, 2014). Fixed parameter models, such as multinomial logit (MNL), ordered logit, and partial proportional odds (PPO), have been frequently employed in pedestrian-injury severity studies. However, the fixed-parameter model neglects the difference across the observations and would result in biased estimations and counter-productive countermeasures (Mannering and Bhat, 2014).

To account for unobserved heterogeneity, random parameter models (or mixed logit models), which allow parameters to vary across observations or groups, have been utilized in many studies (Aziz et al., 2013; Haleem et al., 2015; Kim et al., 2010). Meanwhile, latent-class models (or finite-mixture models) were proposed to address unobserved heterogeneity across groups by segmenting the crash data into homogeneous subsets (Mohamed et al., 2013). Also, a sequential process of combining the latent class model with other discrete outcome models, such as PPO and mixed logit, was employed to further investigate the heterogeneity across pedestrian-injury observations in each group (Behnood and Mannering, 2016; Li et al., 2019a). Recently, the random parameter models are further extended to allow heterogeneity in means and variances by assuming that random parameters are specifically distributed across observations (Behnood and Mannering, 2017). Abay

Table 1
Summary of methods for pedestrian-injury severity studies.

Model	Specific scenario	Data Years	Location	No. of obs.	Authors, Year
Multinomial logit model (MNL)	–	2005–2012	North Carolina	3553	(Chen and Fan, 2019)
Partial proportional odds model (PPO)	Pedestrian age groups	2007–2014	North Carolina	10,875	(Li and Fan, 2019b)
Support vector machine and MNL	Time of day	2010–2014	California	8573	(Mokhtarimousavi, 2019)
Binary logistic regression and tree-based models	–	2014–2016	Changsha, China	791	(Hu et al., 2020)
Classification and regression tree with random forest approach.	Weather	2013	Britain	14,174	(Li et al., 2017)
Extracted rules from Bayesian networks	Urban and suburban	2009–2011	Jordan	21,852	(Mujalli et al., 2019)
Considering unobserved heterogeneity					
Mixed logit model	Signalized and non-signalized locations	2008–2010	Florida	7630	(Haleem et al., 2015)
Mixed logit model	–	1997–2000	North Carolina	5808	(Kim et al., 2010)
Random-parameter (mixed) logit	–	2002–2006	New York City	4666	(Aziz et al., 2013)
Artificial neural network and random parameter ordered response models	Day of week	2010–2014	California	10,146	(Mokhtarimousavi et al., 2020)
Ordered logit, mixed ordered logit, multinomial logit, mixed logit	–	1998–2009	Denmark	4952	(Abay, 2013)
Ordered logit model, generalized ordered logit model, and latent class ordered logit model	–	2002–2006	New York City	4701	(Yasmin et al., 2014)
Latent class clustering and MNL	Whole and each cluster	2006–2015	Louisiana	14,236	(Sun et al., 2019)
Latent class clustering and binary logit	Whole and each cluster	2009–2012	Switzerland	9659	(Sasidharan et al., 2015)
Latent class with ordered probit method, K-means with MNL	Whole and each cluster	2002–2006 (NYC), 2003–2006 (M)	New York City (NYC), Montreal (M)	5820	(Mohamed et al., 2013)
Latent class clustering and PPO	Each cluster	2007–2014	North Carolina	10,875	(Li et al., 2019a)
Latent-class logit and mixed logit models.	Period (pre-recession, recession, and post-recession)	2005–2012	Chicago	19,895	(Behnood and Mannering, 2016)
Considering spatial/temporal patterns					
Bernoulli model and logistic regression	Spatial clusters	2000–2007	Georgia	7763	(Dai, 2012)
Kernel density estimation analysis and MNL	Spatiotemporal pattern	2001–2013	Tunisia	1922	(Ouni and Belloumi, 2018)
Geographically and temporally weighted ordinal logistic regression	Spatiotemporal pattern	2007–2014	North Carolina	13,854	(Liu et al., 2019)
Bayesian spatial Poisson-lognormal model	Signalized intersection	2011–2018	Texas	655	(Munira et al., 2020)

Note: No. of obs. denotes Number of Observations.

(2013) compared the pedestrian-injury severity estimations between multinomial logit, mixed logit, ordered logit, and mixed ordered logit. The results indicated that mixed models outperform fixed models, and fixed models would underestimate the effects of some important behavioral attributes.

For logit models that account for random-effects across observations/groups, using frequentist methods with maximum likelihood estimation (MLE) can be numerically difficult and cumbersome (Chen et al., 2016). In Bayesian models, all estimators are assumed to obey certain distributions. The prior information on the unknown parameter is updated by the likelihood of the observed data. The Markov Chain Monte Carlo (MCMC) simulation method for the inferred posterior distribution could well estimate models with complicated hierarchical structures and capture unobserved heterogeneity across observations (Huang et al., 2008). Several studies pointed out that the Bayesian inference method outperforms the point estimation MLE methods, and posterior probability distributions are ideally suited for analysis with random effects (Chen et al., 2015; Huang et al., 2008; Mitra and Washington, 2007).

Numerous studies adopted hierarchical Bayesian methods in addressing vehicle-related safety issues, while most of them focused on crash frequency outcomes and vehicle-vehicle crashes (Bhat et al., 2017; Li et al., 2019). Munira et al. (2020) employed a Bayesian spatial Poisson-lognormal model to explore the frequency of each pedestrian crash severity at signalized intersections. Also, most previous studies only considered random effects on intercept terms to simplify the model and did not provide the comparison of different hierarchical structures (Chen et al., 2016; Huang et al., 2008; Li et al., 2018). Huang et al. (2008) proposed a hierarchical Bayesian random-intercept-only logit model for driver injury in intersection-related crashes. A two-level structure including observation-level and crash-level was utilized to account for variances within the crashes and between crashes, respectively. As mentioned in (Mannering et al., 2016), many explanatory variables, such as human, vehicle, roadway, traffic, and environment characteristics, would also have heterogeneous effects on the likelihood of the crash injury severity. Hence, an extension of the hierarchical structure that accounts for random-effects on all parameters across observations is supposed to reflect the reality more accurately and further capture the unobserved heterogeneity.

3. Methodology

3.1. Spatiotemporal analysis

3.1.1. Spatiotemporal trend analysis

As shown in Fig. 1, the research area is firstly divided into square bins with a specific distance interval and time interval. Instead of using empirical methods to set the distance interval of the bin and the corresponding neighborhood distances, the average nearest neighbor and the Global Moran's I test are employed, respectively. Then the spatiotemporal analysis is conducted based on these distance values.

To identify the spatially aggregated/dispersed patterns, the Getis-Ord Gi* index (Getis and Ord, 2010) is employed by investigating each feature within the neighboring distance. This index calculates the ratio of the local sum for a feature and its neighbors to the sum of all features. Hence the Gi index could identify the high/low value clusters and outliers. And these spatial patterns are closer to reality patterns compared to the kernel density estimation.

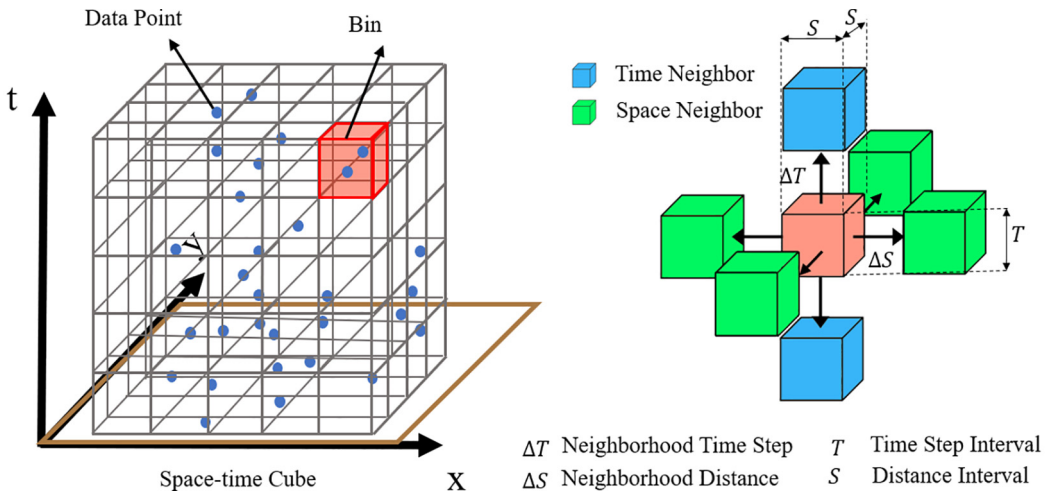

Fig. 1. Illustration of the spatiotemporal trend analysis.

$$G_i^* = \frac{\sum_{j=1}^{n} \omega_{i,j} x_j - \bar{X} \sum_{j=1}^{n} \omega_{i,j}}{SD(x_j) \sqrt{\frac{n}{n-1} \sum_{j=1}^{n} \omega_{i,j}^2 - \frac{1}{n-1} \left(\sum_{j=1}^{n} \omega_{i,j} \right)^2}} \tag{1}$$

where the Getis-Ord Gi* index G_i^* is a statistic Z-score. x_j denotes the attribute value of the jth bin. $\omega_{i,j} = 1$ if the jth bin is within the neighborhood distance interval of the ith bin and 0 otherwise. n is the total number of bins within the neighborhood distance. $\bar{X}$ represents the mean value of x_j. $SD(x_j)$ means the standard deviation for x_j. With statistical significance p-values, $G_i^* > 0$ denotes a clustering of high values, $G_i^* = 0$ means a random pattern of the values, and $G_i^* < 0$ represents a clustering of low values.

Then the Mann-Kendall trend test is performed on Gi* values $\{G_t : t = 1, 2, \cdots, T\}$ at each location/grid with a specified time interval (Kendall and Gibbons, 1990; Mann, 1945). The trend test statistic S is:

$$S = \sum_{i=1}^{T-1} \sum_{j=i+1}^{T} a_{ij} \tag{2}$$

$$a_{ij} = sign(N_j - N_i) = \begin{cases} 1 G_i < G_j \\ 0 G_i = G_j \\ -1 G_i > G_j \end{cases} \tag{3}$$

where a_{ij} is a symbolic variable which denotes the trend/rank of the Gi* values.

The null hypothesis is S = 0, which indicates no trend in the values over time. The Z-score for statistic S is:

$$Z_S = \begin{cases} \frac{S-1}{SD(S)}, S > 0 \\ 0, S = 0 \\ \frac{S+1}{SD(S)}, S < 0 \end{cases} \tag{4}$$

where $SD(S)$ denotes the stand deviation of the S. For a given confidence level α, the null hypothesis is rejected when $|Z_S| \geq |Z_{S,1-\alpha/2}|$. Also, $Z_S > 0$ and $Z_S < 0$ indicate the uptrend and downtrend in bin values, respectively.

3.1.2. Average nearest neighbor for the distance interval

The average nearest neighbor (ANN) calculates the ratio of the average distance of the actual data to the average distance of the data with a hypothetical random distribution (Ebdon, 1985). Therefore, the observed nearest neighbor distance for the clusters could provide a model-based reference value for the distance interval of the bin/grid.

$$ANN = \frac{\bar{D_O}}{\bar{D_E}} = \frac{\sum_{i=1}^{n} d_i / n}{0.5 / \sqrt{n/A}} = \frac{2 \sum_{i=1}^{n} d_i}{\sqrt{nA}} \tag{5}$$

$$Z_{ANN} = \frac{\bar{D_O} - \bar{D_E}}{SD} \tag{6}$$

where $\bar{D_O}$ denotes the observed average distance between each data point and its nearest neighbor. $\bar{D_E}$ represents the expected average distance for the randomly distributed data points. d_i indicates the distance between data i and its nearest neighbor. n denotes the total number of the data points. A indicates the area of a minimum rectangle enclosing all data points. If $ANN1$ = 1, no trend is detected. For ANN<1 and ANN>1, the points exhibit aggregated and dispersed pattern, respectively.

3.1.3. Spatial autocorrelation test for the neighborhood distance

The Global Moran's I, which is a spatial autocorrelation test, evaluates the clustered, dispersed, and random spatial patterns in observations from a global perspective (Moran, 1948). The result of the Global Moran's I index with decreasing weights (weights are proportional to an inverse distance squared function) is utilized to indicate the impact range/radius of the clusters. This range and the number of spatial pattern features identified could further provide a reasonable reference for the neighborhood distance of the bins.

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \times C_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \times D(x_i)} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \times \left(x_i - \bar{x} \right) \left(x - \bar{x} \right)}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \times \frac{1}{n} \sum_{1}^{n} \left(x_i - \bar{x} \right)^2} \tag{7}$$

$$Z_I = \frac{I - E(I)}{\sqrt{D(I)}} \tag{8}$$

where x_i denotes the attribute value of j spatial grid/location. $w_{ij} = 1/\mathrm{d}^2$ if the jth grid is within the spatial neighborhood distance of the ith grid and 0 otherwise, d is the distance between to features. C_{ij} indicates the attribute similarity matrix. $E(I) = -1/(n-1)$, and $D(I) = \mathrm{E}\left(I^2\right) - \mathrm{E}(I)^2$. With statistical significance p-value, 1) if I is positive and close to 1, it represents the incremental spatial autocorrelation (i.e., clustered pattern) within neighborhoods; 2) if I = 0, it indicates a random pattern; and 3) if I < 0, it represents a dispersed pattern.

3.2. Hierarchical Bayesian random-effects model

3.2.1. Random-effects and random-effects-only models

The probability of observation i with severity j in outcome situation t (t = 1, . . ., T) is formulated with a logit link function. When t denotes the outcome set for each observation (i.e., T is the observation numbers), the model accounts for observation-level random-effects. When t denotes the outcome set for each group (i.e., T is the number of the spatiotemporal pattern groups in this paper), the model accounts for group-level random-effects.

$$logit(P_{ijt}) = \log\left(\frac{P_{ijt}}{P_{ij1}}\right) = \boldsymbol{\beta_i'} \boldsymbol{X}_{ijt} + \boldsymbol{\gamma_i'} \boldsymbol{Z}_{ijt} + \varepsilon_{ijt} \tag{9}$$

where P_{ijt} is the probability of the severity j for observation i in outcome situation t. $\boldsymbol{X}_{ijt}$ denotes the fixed variables vector, $\boldsymbol{\beta}_i$ is the corresponding vector of fixed coefficients. $\boldsymbol{Z}_{ijt}$ represents the random variables vector, and $\boldsymbol{\gamma}_i$ is the random coefficients vector corresponding to $\boldsymbol{Z}_{ijt}$.

It is assumed that each $\boldsymbol{\gamma}_i$ is drawn from a normally distributed super-population. A revised hierarchical structure is added to the model in which a prior for the covariance matrix Ω_γ is specified:

$$\pi(\boldsymbol{\gamma}_i) = N(0, \boldsymbol{\Omega}_\gamma) \tag{10}$$

$$\pi(\boldsymbol{\Omega}_\gamma) = inverse\ Wishart(v_0, \boldsymbol{V}_0) \tag{11}$$

The conjugate prior of the covariance matrix Ω_γ obeys an inverse Wishart distribution. v_0 is the degrees of freedom and $\boldsymbol{V}_0$ is the identity matrix for all random parameters. The Ω_γ characterizes the extent of heterogeneity among observations or groups. Large diagonal elements of Ω_γ represent the heterogeneity in part-worths. Off-diagonal elements indicate patterns in the evaluation of attributes in pairs.

Some studies assume that the unobserved heterogeneity is inherent in the intercepts or individual characteristic variables (Huang et al., 2008; Rossi et al., 2012). To make comparisons between these models, a random intercept model and a model with random-effects on both intercepts and individual characteristics are proposed. To further account for heterogeneity, this paper adopts a random-effects-only model with a revised hierarchical structure that includes no fixed-effects ($\boldsymbol{\beta}_i = 0$.) but only random-effects (Rossi et al., 2012).

$$\pi(\boldsymbol{\gamma}_i) = N(\overline{\boldsymbol{\gamma}}_i, \boldsymbol{\Omega}_\gamma) \tag{12}$$

where $\overline{\boldsymbol{\gamma}}_i$ is the mean vector of random coefficients, and it represents the average utility (or part-worths) across the respondents (or injury severity levels).

In this study, Bayesian non-informative priors with normal distributions (0,1000) are applied to infer the unknown parameters as referred by (Chen et al., 2016; Huang et al., 2008). A Markov Chain Monte Carlo (MCMC) method with Gamerman Metropolis sampling is applied for model estimation. The chain is simulated with a "burn-in" part and a simulation part. Also, a thinning process, which samples the posterior distribution data with a certain interval, is used to reduce autocorrelation of the estimators. The trace plots, posterior autocorrelations, effective sample sizes, and Monte Carlo standard errors are utilized to ensure that "burn-in" iterations are sufficiently simulated and the results reach convergences. A variable is considered to be significant when 95% Bayesian credible interval (BCI) of its posterior mean does not cover 0 and not significant otherwise (Li et al., 2018).

3.2.2. Performance measurement

For performance measurement and selection of the Bayesian models, Deviance Information Criterion (DIC) was proposed by Spiegelhalter et al. (2002). The DIC is generalized by two hierarchical modeling measurements: Akaike information criterion (AIC) and Bayesian information criterion (BIC).

$$DIC = D\left(\bar{\gamma}\right) + 2pD = \bar{D(\gamma)} + pD \tag{13}$$

where, $D\left(\bar{\gamma}\right)$ denotes the deviance evaluated at the posterior means of the parameters γ and is taken as a measure of the model suitability. $\bar{D(\gamma)}$ represents the posterior means of the deviance of the estimated parameters γ and is treated as a

measure of model fit. *pD* is the effective number of parameters and is considered as a measure of model complexity, $pD = \bar{D(\gamma)} - D\left(\bar{\gamma}\right)$. For model comparison, a lower DIC value indicates a superior model for parameter estimation and model fit.

The hit probability calculates the average of estimated probabilities of outcome correctly in the input dataset. And the hit probability is used to measure the goodness-of-fit of the discrete outcome model.

$$hitrate = \frac{1}{n}\sum\nolimits_{i}^{n} P_{ij}(ifpredictchoicej = correctchoice) \tag{14}$$

3.2.3. Marginal effect

To interpret the results of random-effects models with category variables (dummied with 1 to denote the presence of the variable and 0 otherwise), marginal effects are utilized to illustrate the impact of the explanatory variable in the changing values of severity probability outcomes (Song and Fan, 2020).

$$E_{X_{ij}}^{P_{ij}} = \frac{1}{n}\sum_{i=1}^{n}\left[P_{ij}(X_{ij}=1) - P_{ij}(X_{ij}=0)\right] \tag{15}$$

where the average difference value of P_{ij} over all observations is calculated when the dummied explanatory variable X_{ij} changes from 0 to 1.

4. Data description

A total number of 33,707 police-reported pedestrian-vehicle crash observations in North Carolina from 2007 to 2018 are collected from the North Carolina Department of Transportation (NCDOT). To conduct the spatiotemporal analysis, the distance interval is set as 382 m based on the ANN test result (ANN ratio: 0.286; Z-score: −250.845; P-value <0.0001). The time interval is set as 1 year to investigate the annually temporal patterns and mitigate the seasonal variations of the crash data. As shown in Fig. 2, an 8000 m neighborhood distance is identified based on the results of the spatial autocorrelation test and the total number of the spatiotemporal locations. The corresponding Moran's Index is 0.36 (z-score = 322, P-value <0.0001), which indicates an aggregated spatial pattern within the neighborhood distance. Fig. 3 illustrates the spatiotemporal pattern results of the pedestrian-injury crash locations in North Carolina. Also, Fig. 3 shows an upward aggregated tendency of the crash locations in urban areas of Charlotte.

To model the pedestrian-injury severities, a total number of 27,091 observations are utilized after selecting the pedestrian with the highest injury severity in single-vehicle involved crashes and deleting observations with missing values. A detailed summary of ten identified spatiotemporal patterns and the corresponding statistics of the pedestrian-injury severities are presented in Table 2. Considering the spatial patterns of the crashes, the numbers of the crashes with aggregated, dispersed, and no-trend patterns are 12,890, 230, and 13,970, respectively. Meanwhile, six temporal patterns are identified including new, consecutive, intensifying, persistent, sporadic, and historical patterns. The pedestrian-injury severity is classified into three levels (i.e., severe injury [including fatal/incapacitating injury], minor injury [including non-incapacitating injury], and no injury [including no/possible injury]). Table 3 exhibits the statistics of explanatory variables of the pedestrian-injury severities. The explanatory variables are classified into categories of pedestrian, vehicle, crash, locality and roadway, environment and time, and traffic control.

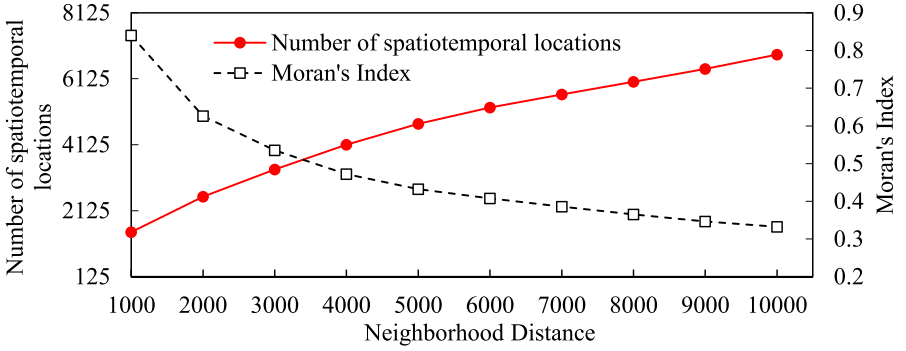

Fig. 2. Results of the Moran's Index and the total number of spatiotemporal locations identified under different neighborhood distance.

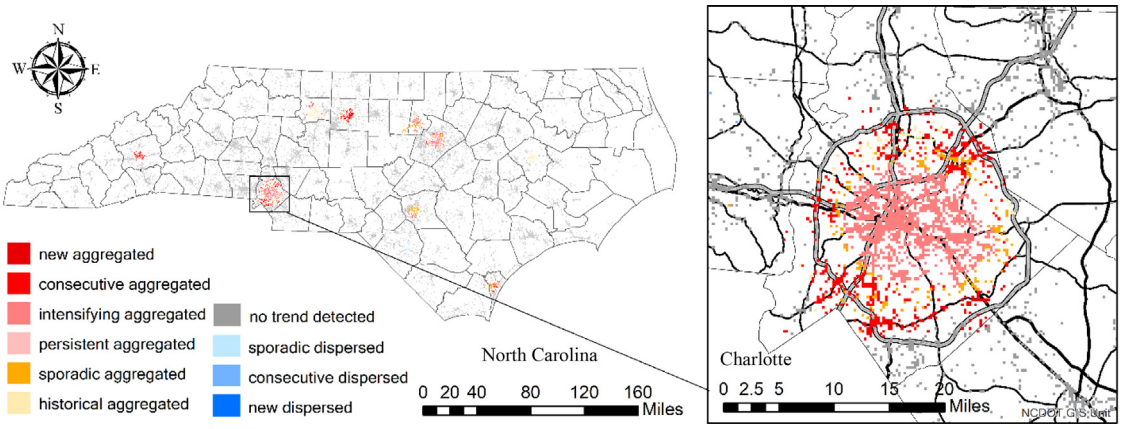

Fig. 3. Spatiotemporal patterns of the pedestrian-injury crash locations in North Carolina.

Table 2
Description and statistics for spatiotemporal patterns of pedestrian-injury crashes.

NO	Spatiotemporal patterns	Description	Total	Severe injury	Minor injury	No injury
1	New aggregated	A location with a statistically significant spatially aggregated pattern for the final time step and has never been a statistically significant aggregated pattern before.	432	56 (12.96%)	156 (36.11%)	220 (50.93%)
2	Consecutive aggregated	A location with a single uninterrupted run of statistically significant aggregated patterns to the final time-step interval. The location has never been a statistically significant aggregated pattern prior to the final uninterrupted run of aggregated patterns. Also, less than 90% of all bins are statistically significant aggregated patterns.	3051	322 (10.55%)	1085 (35.56%)	1644 (53.88%)
3	Intensifying aggregated	A location that has been a statistically significant aggregated pattern for 90% of the time-step intervals, including the final time step. Also, the aggregated patterns have a significant intensifying pattern over time.	5699	574 (10.07%)	2166 (38.01%)	2959 (51.92%)
4	Persistent aggregated	A location that has been a significant aggregated pattern for 90% of the time-step intervals. The aggregated patterns have no trend over time.	54	7 (12.96%)	22 (40.74%)	25 (46.3%)
5	Sporadic aggregated	Less than 90% of the time-step intervals have been statistically significant aggregated patterns and none of the time-step intervals have been statistically significant dispersed patterns.	2533	310 (12.24%)	957 (37.78%)	1266 (49.98%)
6	Historical aggregated	The most recent periods are not aggregated pattern, but at least 90% of the time-step intervals have been statistically significant aggregated pattern.	1121	164 (14.63%)	443 (39.52%)	514 (45.85%)
7	No trend detected	Does not fall into any of the aggregated or dispersed patterns.	1,3971	2609 (18.67%)	5163 (36.96%)	6199 (44.37%)
8	New dispersed	A location with a statistically significant dispersed pattern for the final time step and has never been a statistically significant dispersed pattern before.	17	7 (41.18%)	6 (35.29%)	4 (23.53%)
9	Consecutive dispersed	A location with a single uninterrupted run of statistically significant dispersed patterns to the final time-step interval. The location has never been a statistically significant dispersed pattern prior to the final uninterrupted run of dispersed patterns. Also, less than 90% of all bins are statistically significant dispersed patterns.	31	3 (9.68%)	11 (35.48%)	17 (54.84%)
10	Sporadic dispersed	Less than 90% of the time-step intervals have been statistically significant dispersed patterns and none of the time-step intervals have been statistically significant aggregated patterns.	182	30 (16.48%)	56 (30.77%)	96 (52.75%)

5. Likelihood ratio tests

To statistically determine if the temporal instability is significant during the study period, a series of likelihood ratio tests are conducted according to (Washington et al., 2011).

$$X^2 = -2\left[LL(\boldsymbol{\beta}_{t_2 t_1}) - LL(\boldsymbol{\beta}_{t_1})\right] \tag{16}$$

where $LL(\boldsymbol{\beta}_{t_2 t_1})$ is the log-likelihood at the convergence of a model using the converged parameters from time t_2 (with restricting the parameters to be the estimated parameters of time t_2), while using data from time t_1. $LL(\boldsymbol{\beta}_{t_1})$ is the log-likelihood at the convergence of the model with data of time t_1. This test is also reversed by using $LL(\boldsymbol{\beta}_{t_1 t_2})$ and $LL(\boldsymbol{\beta}_{t_2})$. The resulting value X^2 is χ^2 distributed with degrees of freedom being equal to the number of estimated parameters in $\boldsymbol{\beta}_{t_1 t_2}$. The null hypothesis is that the parameters in time t_1 and t_2 are equal.

Table 3
Statistics of explanatory variables for pedestrian-injury severities during single pedestrian-vehicle crashes.

Variable		Description	Total	Pedestrian-Injury Severity		
				Severe injury	Minor injury	No injury
		Number of observations	27,091	4082(15.07%)	10065(37.15%)	12944(47.78%)
Pedestrian Characteristics						
PedAge	1	PedAge $\leq$ 24	8584	1077(12.55%)	3520(41.01%)	3987(46.45%)
	2	24 < PedAge $\leq$ 54	13,017	2038(15.66%)	4558(35.02%)	6421(49.33%)
	3	55 < PedAge $\leq$ 64	2663	474(17.8%)	896(33.65%)	1293(48.55%)
	4	PedAge $\geq$ 65	2827	493(17.44%)	1091(38.59%)	1243(43.97%)
PedSex	1	Male	16,036	2843(17.73%)	6093(38%)	7100(44.28%)
	2	Female	11,055	1239(11.21%)	3972(35.93%)	5844(52.86%)
PedAlcFlag	1	PedAlcFlag = 'No'	22,740	2532(11.13%)	8455(37.18%)	11753(51.68%)
	2	PedAlcFlag = 'Yes'	3119	990(31.74%)	1256(40.27%)	873(27.99%)
	3	Unknow	1232	560(45.45%)	354(28.73%)	318(25.81%)
Vehicle Type						
DrvrVehTyp	1	Small	15,144	2002(13.22%)	5694(37.6%)	7448(49.18%)
	2	Middle	11,165	1834(16.43%)	4105(36.77%)	5226(46.81%)
	3	Heavy	782	246(31.46%)	266(34.02%)	270(34.53%)
Crash Characteristics						
CrashGrp	1	Walking Along Roadway	2727	522(19.14%)	1021(37.44%)	1184(43.42%)
	2	Crossing Roadway with Vehicle Not Turning	5078	1281(25.23%)	1943(38.26%)	1854(36.51%)
	3	Crossing Roadway with Vehicle Turning	2978	111(3.73%)	1041(34.96%)	1826(61.32%)
	4	Off Roadway	3337	164(4.91%)	1046(31.35%)	2127(63.74%)
	5	Pedestrian in Roadway	1944	569(29.27%)	649(33.38%)	726(37.35%)
	6	Dash/Dart-Out	2264	414(18.29%)	1150(50.8%)	700(30.92%)
	7	Backing Vehicle	2737	126(4.6%)	771(28.17%)	1840(67.23%)
	8	Multiple Threat/Trapped	309	28(9.06%)	148(47.9%)	133(43.04%)
	9	Bus related Vehicle	220	35(15.91%)	97(44.09%)	88(40%)
	10	Other/Unusual Circumstances	5497	832(15.14%)	2199(40%)	2466(44.86%)
DriAlcFlag	1	DriAlcFlag = 'No'	23,101	3464(15%)	8607(37.26%)	11030(47.75%)
	2	DriAlcFlag = 'Yes'	981	298(30.38%)	377(38.43%)	306(31.19%)
	3	Unknow	3009	320(10.63%)	1081(35.93%)	1608(53.44%)
AmbulanceR	1	Ambulance Rescue	6758	350(5.18%)	1609(23.81%)	4799(71.01%)
	2	No Ambulance Rescue	20,333	3732(18.35%)	8456(41.59%)	8145(40.06%)
HitRun	1	No Hit and Run	24,232	3754(15.49%)	9094(37.53%)	11384(46.98%)
	2	Hit and Run	2859	328(11.47%)	971(33.96%)	1560(54.56%)
Locality and Roadway Characteristics						
Locality	1	Rural	7379	1841(24.95%)	2780(37.67%)	2758(37.38%)
	2	Urban	19,712	2241(11.37%)	7285(36.96%)	10186(51.67%)
Development	1	Residential	9323	1319(14.15%)	3810(40.87%)	4194(44.99%)
	2	Commercial	13,837	1733(12.52%)	4833(34.93%)	7271(52.55%)
	3	Industrial	147	21(14.29%)	58(39.46%)	68(46.26%)
	4	Institutional	915	52(5.68%)	327(35.74%)	536(58.58%)
	5	Farms, Woods, Pastures	2869	957(33.36%)	1037(36.14%)	875(30.5%)
RdCurve	1	straight	25,667	3705(14.43%)	9527(37.12%)	12435(48.45%)
	2	Curve	1424	377(26.47%)	538(37.78%)	509(35.74%)
RdGrad	1	Level	22,004	3080(14%)	8092(36.78%)	10832(49.23%)
	2	Grade	3923	802(20.44%)	1497(38.16%)	1624(41.4%)
	3	Hillcrest	966	163(16.87%)	379(39.23%)	424(43.89%)
	4	Bottom	198	37(18.69%)	97(48.99%)	64(32.32%)
RdClass	1	US Route	1643	592(36.03%)	577(35.12%)	474(28.85%)
	2	Interstate	510	259(50.78%)	141(27.65%)	110(21.57%)
	3	State Route	1454	440(30.26%)	525(36.11%)	489(33.63%)
	4	State Secondary Route	2826	721(25.51%)	1132(40.06%)	973(34.43%)
	5	Local Street, Driveway	14,026	1782(12.7%)	5715(40.75%)	6529(46.55%)
	6	Public Vehicular Area	6632	288(4.34%)	1975(29.78%)	4369(65.88%)
RdConfig	1	One-Way, Not Divided	2315	166(7.17%)	818(35.33%)	1331(57.49%)
	2	Two-Way, Not Divided	19,808	2744(13.85%)	7275(36.73%)	9789(49.42%)
	3	Two-Way, Divided	4968	1172(23.59%)	1972(39.69%)	1824(36.71%)
Environment and time Characteristics						
LightCond	1	Daylight	15,344	1321(8.61%)	5651(36.83%)	8372(54.56%)
	2	Dawn/Dusk Light	1170	155(13.25%)	424(36.24%)	591(50.51%)
	3	Dark – Lighted Roadway	5796	1006(17.36%)	2291(39.53%)	2499(43.12%)
	4	Dark – Roadway Not Lighted	4781	1600(33.47%)	1699(35.54%)	1482(31%)

(*continued on next page*)

Table 3 (*continued*)

Variable		Description	Total	Pedestrian-Injury Severity		
				Severe injury	Minor injury	No injury
Weather	1	Clear	20,886	3092(14.8%)	7800(37.35%)	9994(47.85%)
	2	Cloudy	3782	622(16.45%)	1362(36.01%)	1798(47.54%)
	3	Rain	2121	313(14.76%)	784(36.96%)	1024(48.28%)
	4	Snow, Sleet, Hail, Freezing Rain/Drizzle	180	17(9.44%)	77(42.78%)	86(47.78%)
	5	Fog, Smog, Smoke	122	38(31.15%)	42(34.43%)	42(34.43%)
Hour	1	10:00–12:59	3452	266(7.71%)	1155(33.46%)	2031(58.84%)
	2	13:00–16:59	6313	531(8.41%)	2267(35.91%)	3515(55.68%)
	3	17:00–21:59	9497	1619(17.05%)	3692(38.88%)	4186(44.08%)
	4	22:00–5:59	4288	1209(28.19%)	1623(37.85%)	1456(33.96%)
	5	6:00–9:59	3541	457(12.91%)	1328(37.5%)	1756(49.59%)
Traffic Control Types						
Traffic control	1	No Control Present	18,220	2681(14.71%)	6737(36.98%)	8802(48.31%)
	2	Signs	1948	165(8.47%)	682(35.01%)	1101(56.52%)
	3	Signal	3815	407(10.67%)	1502(39.37%)	1906(49.96%)
	4	Double Yellow Line, No Passing Zone	2814	809(28.75%)	1056(37.53%)	949(33.72%)
	5	Human Control	294	20(6.8%)	88(29.93%)	186(63.27%)

Note: variables numbered with 1 are set as the base for the explanatory variables.

Table 4 presents the results of the likelihood ratio tests between every two years based on the multinomial logit models, which are utilized as the basic models for variable selections (Greene et al., 2005). It is showed that the statistics for some periods show a probability of less than 95% to reject the null hypothesis that the parameters of the two periods are equal. However, the results still indicate the existence of the temporal instability of the pedestrian-vehicle crash factors as the reversed tests for those periods all present a probability of more than 99% to reject the null hypothesis. As mentioned in (Behnood and Mannering, 2016), the inherent reasons for the temporal instability of the factors for pedestrian-vehicle crashes are yet to be determined. The investigating of the spatiotemporal patterns of the pedestrian-vehicle crashes might provide insights into possible inherent reasons for the spatiotemporal instability of the factors and crashes

Meanwhile, to test for the distinction between factors of different spatiotemporal pattern crashes, the following likelihood ratio tests are also applied according to (Washington et al., 2011).

$$X^2 = -2\left[LL(\boldsymbol{\beta}_{total}) - \sum\nolimits_i^n LL(\boldsymbol{\beta}_{subset\ i})\right] \tag{17}$$

where $LL(\boldsymbol{\beta}_{total})$ is the log-likelihood at the convergence of a model containing the converged parameters based on the total data. $LL(\boldsymbol{\beta}_{subset\ i})$ denotes the log-likelihood at the convergence of a model containing the converged parameters based on the subset *i*. The degrees of freedom are calculated by the summation of the number of estimated parameters in all subset models minus the number of estimated parameters in the whole dataset model. The X^2 is χ^2 distributed with the null hypothesis that the parameters for the segmented subsets are equal.

In this paper, the whole dataset is further segmented by the spatiotemporal patterns and the log-likelihood values are estimated by the basic multinomial logit models. For subsets with new aggregated, consecutive aggregated, intensifying aggregated, sporadic aggregated, and historical aggregated patterns, the value of X^2 is 154.6 which is χ^2 distributed with 102 degrees of freedom. This X^2 value gives 99.93% confidence to reject the null hypothesis that the parameters for these spatiotemporal pattern subsets are the same. This result indicates the significant distinctions between the factors of the crashes with different spatiotemporal patterns. According to (Behnood and Mannering, 2016), shifts in pedestrian/driver behaviors and changes in economic conditions are all possible reasons for the temporal instability of the factors of the crash

Table 4
Likelihood ratio test results between different period pairs (χ2 values with the degrees of freedom in brackets and the confidence level in parenthesis).

t1	t2					
	2007–2008	2009–2010	2011–2012	2013–2014	2015–2016	2017–2018
2007–2008	–	98 [44] (>99.99%)	52 [41] (>88.35%)	50 [39] (>88.84%)	22 [45] (>0.15%)	34 [44] (>13.85%)
2009–2010	138 [37] (>99.99%)	–	70 [42] (>99.57%)	82 [41] (>99.98%)	62 [47] (>92.99%)	74 [45] (>99.58%)
2011–2012	214 [37] (>99.99%)	72 [45] (>99.35%)	–	60 [41] (>97.21%)	42 [47] (>32.07%)	54 [45] (>83.18%)
2013–2014	180 [37] (>99.99%)	82 [45] (>99.94%)	60 [42] (>96.47%)	–	22 [47] (>0.06%)	50 [45] (>71.85%)
2015–2016	260 [37] (>99.99%)	144 [45] (>99.99%)	126 [42] (>99.99%)	82 [41] (>99.98%)	–	102 [45] (>99.99%)
2017–2018	302 [37] (>99.99%)	126 [45] (>99.99%)	110 [42] (>99.99%)	100 [41] (>99.99%)	98 [47] (>99.99%)	–

injury severity. Therefore, this paper further explores the factors that contribute to the pedestrian-injury severity in crashes with different spatiotemporal patterns.

6. Model comparisons

This paper proposes two ways to account for spatiotemporal features of the pedestrian crash in hierarchical Bayesian models, and they are: 1) a model with random-effects across spatiotemporal segmented groups, and 2) models with random-effects across observations in each spatiotemporal segmented data. As shown in Table 5, the basic Bayesian multinomial logit (BMNL) model and three variants of the hierarchical Bayesian random-effects logit (HBREL) models are employed for the whole dataset. First, a hierarchical Bayesian random-intercept logit model with random-effects across groups (HBRIL-grp) is established. The HBRIL-grp model could account for the heterogeneity across the spatiotemporal groups with the whole dataset. In this case, different spatiotemporal patterns of the crash data could be addressed within one model simultaneously. Also, this model shows slight improvements in the DIC (smaller value of DIC indicate a better goodness-of-fit) and the hit probability (which denotes the prediction accuracy of the model) compared to the BMNL model. To capture the heterogeneity across the observations, a HBRIL model with random-effects across observations (HBRIL-obs.) and a HBREL model with random-effects on intercepts and individual characteristics across observations (HBREL-inter. and indiv.) are developed. The results of two models indicate that the increase of random parameters and accounting heterogeneity across observations could significantly improve the performance of the DIC (drops from 47,115 to 45,042 and 17,490, respectively) and hit probability (increases from 47% to 69.8% and 99.9%, respectively). However, these two models could not capture the heterogeneity across different spatiotemporal groups.

To further account for the unobserved heterogeneity in the segmented spatiotemporal pattern dataset, hierarchical Bayesian random-effects-only logit (HBREOL) models with random-effects across observations are further utilized. All parameters in HBREOL models are treated as random parameters. It is noted that the numbers of observations in dispersed spatial patterns and the persistent aggregated pattern are too less to establish separate models. Hence, data with the aggregated spatial patterns, which denote crashes occurred in crash-prone areas, are utilized considering different temporal patterns. Table 6 shows the results of HBREOL models for new aggregated, consecutive aggregated, intensifying aggregated, sporadic aggregated, and historical aggregated datasets. The HBREOL models significantly decrease the DIC and increase the hit probability to 99% compared to the BMNL model.

7. Model results and discussion

7.1. Parameter estimation results

7.1.1. Parameter estimations for the whole dataset model

For the whole dataset model, a hierarchical Bayesian random-intercept logit model with random-effects across groups (HBRIL-grp) is established. Table 7 shows the coefficient estimation results and the simulation setups of the model. The HBRIL-grp model captures the heterogeneity by allowing the intercept terms to vary across the spatiotemporal pattern groups. Table 8 indicates the specific estimation results of the intercepts for each group and the covariance of the random intercept terms. There is a small difference between the intercept terms of 10 spatiotemporal pattern groups. For crashes with the spatial aggregated pattern, persistent aggregated crashes show a larger utility in severe and minor injuries. This indicates that the crashes with persistent aggregated patterns would result in more severe injuries than other spatial aggregated crashes.

Table 5
Model comparison of the hierarchical models with the whole dataset.

Dataset	Whole Dataset			
Model	BMNL [a]	BHRIL-grp. [b]	BHRIL-obs. [c]	BHREL-inter. and indiv. [d]
$\overline{D(\gamma)}$(Posterior Mean of Deviance)	47048.3	47003.6	33098.8	8784.6
$D(\overline{\gamma})$(Deviance Evaluated at Posterior Mean)	46981.2	46920.6	21,156	79.2
pD (Effective Number of Parameters)	67.1	82.9	11942.7	8705.3
DIC (Deviance Information Criterion)	47115.4	47086.5	45041.5	17489.9
Hit Probability (Prediction Accuracy)	0.47	0.471	0.698	0.999

Note:
[a] BMNL denotes the Bayesian multinomial logit model;
[b] HBRIL-grp. denotes the hierarchical Bayesian random-intercept logit model with random-effects across spatiotemporal groups;
[c] HBRIL-obs. inter. denotes the hierarchical Bayesian random-intercept logit model with random-effects across observation;
[d] HBREL-inter. and indiv. denotes the hierarchical Bayesian random-effects logit model with random-effects on intercept and individual characteristics (random-effects across observations).

Table 6
Model comparison of the models with the spatiotemporal pedestrian-injury crash data.

Dataset	New aggregated		Consecutive aggregated		Intensifying aggregated		Sporadic aggregated		Historical aggregated	
Model	BMNL[a]	HBREOL[b]	BMNL	HBREOL	BMNL	HBREOL	BMNL	HBREOL	BMNL	HBREOL
$\overline{D(\gamma)}$	719	245	4937.3	1166.8	9406.1	2082.7	4240.4	1004.7	1929.6	465.2
$D(\bar{\gamma})$	704.8	158.8	4897.6	455	9363	390.7	4201	456.9	1903	209.2
pD	14.2	86.2	39.7	711.8	43.1	1692	39.4	547.7	26.6	256
DIC	733.2	331.2	4977.1	1878.5	9449.1	3774.7	4279.8	1552.4	1956.2	721.1
Hit Probability	0.505	0.968	0.508	0.989	0.495	0.991	0.49	0.989	0.487	0.988

Note:
[a] BMNL denotes the Bayesian multinomial logit model.
[b] HBREOL denotes the hierarchical Bayesian random-effects-only logit model (random-effects across observations).

7.1.2. Parameter estimations for segmented spatiotemporal pattern data models

For the dataset segmented by spatiotemporal patterns, the hierarchical Bayesian random-effects-only logit (HBREOL) models are utilized. Tables 9–13 present coefficient estimation results and the simulation setups of the models with new aggregated, consecutive aggregated, intensifying aggregated, sporadic aggregated, and historical aggregated pattern data.

7.2. Marginal effect results

Table 14 presents the results of the marginal effects of the significant variables for the whole dataset model and each spatiotemporal segmented dataset model. The following sections provide a specific analysis of the heterogeneity results based on marginal effects, especially for the impacts of factors on the severe injury.

7.3. Discussions of the significant factors

7.3.1. Human characteristics

The age of the pedestrian is divided into four categories considering different physical conditions and the corresponding proportion of the injury severities. A similar classification was also employed in (Kim et al., 2010). Compared to young pedestrians (age less than 25), elder pedestrians (age larger than 54) increase the probability of severe injury (SI). Especially for crashes with new aggregated pattern, elder pedestrians would increase the probability of the SI by 0.091. However, for middle-age pedestrians (age within 25 and 54), heterogeneous results for the SI could be found in new aggregated crashes (−0.028), historical aggregated crashes (−0.008), and sporadic aggregated crashes (0.009). Yasmin et al. (2014) also found that middle-age pedestrians would decrease injury severities while young and elder pedestrians suffer severer injuries in the crashes. Such heterogeneity effects within different age groups could also be supported by some previous studies (Abay, 2013; Aziz et al., 2013). In comparison with male pedestrians, female pedestrians could slightly decrease the probability of the SI by 0.008 and 0.021 in the whole dataset model and the consecutive aggregated crash model, respectively. A possible reason for this is that female pedestrians are more cautious than male pedestrians and less likely to be involved in severe crashes.

Alcohol-involved pedestrians would increase the probability of the SI (up to 0.067 in intensifying aggregated crashes) in all cases except historical aggregated crashes. Also, alcohol-involved drivers increase the probability of the SI in the whole dataset model (0.076), intensifying aggregated crash model (0.057), and sporadic aggregated crash model (0.033). The increasing effects of the intoxicated pedestrian/driver on the SI in locations with intensifying aggregated crashes might require a more frequent and strict alcohol test in these intensifying crash-prone locations.

7.3.2. Vehicle characteristics

The vehicles are classified into three weight categories as the crash severity is highly positively correlated to the vehicle weight (Aziz et al., 2013). Compared to the crashes that involved small vehicles, pedestrians are more likely to suffer SI with an increase in vehicle weights. For example, crashes involved with heavy and middle size vehicle in locations with intensifying aggregated crashes could increase the probability of the SI by 0.047 and 0.013, respectively. It is suggested to lower speed limits and restrict the permit traveling time for heavy vehicles to avoid the traveling peak hours of the pedestrians in these locations.

7.3.3. Crash characteristics

Compared to crashes with ambulance rescue, crashes without ambulance rescue would increase the probability of the SI in all models (up to 0.098 in the whole dataset model). In locations with new aggregated crashes, crashes without ambulance rescue could increase the probability of the SI to 0.028. Specific attention should be paid to no ambulance rescue crashes in these locations with emerging aggregated crashes. Compared to crashes without hit-and-run, crashes with hit-and-run increase the probability of the SI in the whole dataset model (0.007) and the intensifying aggregated crashes model

Table 7
Results for the hierarchical Bayesian random-intercept logit model with random-effects across groups (spatiotemporal patterns) for the whole dataset.

Variable	Description	Severe Injury				Minor Injury			
		Mean	S.D.	t value	95% BCI	Mean	S.D.	t value	95% BCI
REMean intercept	Radom effects mean of the intercept	−3.182	0.303	−10.5	(−3.78, −2.593)	−0.878	0.241	−3.64	(−1.357, −0.419)
PAge2	24 < PedAge ≤ 54 (PedAge ≤ 24)					−0.122	0.029	−4.21	(−0.122,0.029)
PAge3	55 < PedAge ≤ 64					0.492	0.063	7.81	(0.373,0.62)
PAge4	PedAge ≥ 65	0.975	0.068	14.34	(0.839, 1.106)	0.324	0.051	6.35	(0.233, 0.428)
PAlc2	PedAlcFlag = 'Yes' (No)	0.73	0.062	11.77	(0.614, 0.852)	0.356	0.051	6.98	(0.256, 0.458)
PAlc3	Unknow	1.659	0.088	18.85	(1.487, 1.83)	0.302	0.083	3.64	(0.14, 0.461)
PSex2	PedSex = Female (Male)	−0.113	0.045	−2.51	(−0.194, −0.018)	−0.071	0.029	−2.45	(−0.128, −0.014)
Veh2	Middle (Small)	0.239	0.04	5.98	(0.157, 0.313)				
Veh3	Heavy	1.288	0.114	11.3	(1.071, 1.507)	0.346	0.093	3.72	(0.162, 0.525)
Amb2	No ambulance (Yes)	1.607	0.064	25.11	(1.492, 1.741)	1.041	0.033	31.55	(0.978, 1.107)
Run2	Hit and run (No)					−0.121	0.045	−2.69	(−0.205, −0.028)
DAlc2	DriAlcFlag = 'Yes' (No)	0.785	0.096	8.18	(0.597, 0.968)	0.263	0.083	3.17	(0.107, 0.427)
DAlc3	Unknow	−0.349	0.072	−4.85	(−0.492, −0.213)				
Grp2	Crossing roadway with vehicle not turning (Walking along roadway)	0.886	0.069	12.84	(0.752, 1.025)	0.249	0.047	5.3	(0.163, 0.346)
Grp3	Crossing roadway with vehicle turning	−0.908	0.121	−7.5	(−1.145, −0.672)	−0.229	0.054	−4.24	(−0.334, −0.128)
Grp5	Pedestrian in roadway	0.639	0.073	8.75	(0.496, 0.779)				
Grp6	Dash/dart-out	0.973	0.089	10.93	(0.813, 1.157)	0.665	0.061	10.9	(0.552, 0.784)
Grp7	Backing vehicle	−0.404	0.111	−3.64	(−0.637, −0.201)	−0.249	0.055	−4.53	(−0.351, −0.135)
Grp8	Multiple threat/trapped					0.326	0.123	2.65	(0.099, 0.573)
Grp9	Bus related vehicle	0.799	0.229	3.49	(0.35, 1.241)	0.362	0.155	2.34	(0.07, 0.666)
Grp10	Other/unusual condition	0.437	0.067	6.52	(0.308, 0.564)	0.248	0.041	6.05	(0.169, 0.327)
Urb2	Urban (Rural)	−0.474	0.065	−7.29	(−0.606, −0.355)	−0.25	0.04	−6.25	(−0.332, −0.175)
Deve2	Commercial (Residential)					−0.126	0.03	−4.2	(−0.179, −0.061)
Deve5	Farms, woods, pastures	0.173	0.06	2.88	(0.068, 0.303)				
Curv2	Curve (Straight)	0.361	0.073	4.95	(0.222, 0.503)				
Grad2	Grade (Level)	0.388	0.056	6.93	(0.278, 0.498)	0.105	0.041	2.56	(0.026, 0.185)
Grad4	Bottom					0.435	0.15	2.9	(0.15, 0.736)
Clas2	Interstate (US route)	0.635	0.115	5.52	(0.401, 0.848)				
Clas4	State secondary route	−0.255	0.065	−3.92	(−0.381, −0.134)				
Clas5	Local street, driveway	−0.434	0.063	−6.89	(−0.559, −0.31)				
Clas6	Public vehicular area	−1.307	0.093	−14.05	(−1.482, −1.13)	−0.387	0.04	−9.68	(−0.464, −0.306)
Conf2	Two-way, not divided (One-way, not divided)	0.251	0.094	2.67	(0.052, 0.417)				
Conf3	Two-way, divided	0.635	0.102	6.23	(0.436, 0.835)	0.208	0.041	5.07	(0.134, 0.29)
Ligh2	Dawn/dusk light (Daylight)					−0.177	0.066	−2.68	(−0.316, −0.057)
Ligh3	Dark – lighted roadway	0.349	0.067	5.21	(0.222, 0.483)				
Ligh4	Dark – roadway not lighted	0.583	0.067	8.7	(0.457, 0.716)				
Weat4	Snow, sleet, hail, freezing rain/drizzle (Clear)	−1.126	0.283	−3.98	(−1.688, −0.582)				
Hour3	17:00–21:59 (10:00–12:59)	0.155	0.067	2.31	(0.029, 0.29)	0.174	0.034	5.12	(0.107, 0.242)
Hour4	22:00–5:59	0.608	0.084	7.24	(0.451, 0.774)	0.387	0.047	8.23	(0.298, 0.477)
Hour5	6:00–9:59	0.322	0.075	4.29	(0.176, 0.47)	0.152	0.045	3.38	(0.065, 0.239)
Cntrl2	Signs (No control)	−0.483	0.098	−4.93	(−0.686, −0.305)	−0.251	0.055	−4.56	(−0.354, −0.14)
Cntrl3	Signal	−0.17	0.068	−2.5	(−0.292, −0.031)				
Cntrl4	Double yellow line, no Passing zone					−0.102	0.05	−2.04	(−0.202, −0.011)
Cntrl5	Human control	−1.055	0.254	−4.15	(−1.535, −0.557)	−0.356	0.135	−2.64	(−0.621, −0.098)

Note: The variable in the parentheses is the base category. REMean denotes posterior mean of the random-effects parameter. Observation number: 27,091; Burn-In Size: 50,000; Simulation Size: 15,000; Thinning: 5; DIC: 47086.5; Hit Probability: 0.471.

(0.035). Ma et al. (2018) also found hit-and-run would increase the injury severity of pedestrians and observed an increasing trend of hit-and-run behaviors in young and elder drivers. The increasing tendency of the hit-and-run crash requires one to increase the frequency of patrols or install surveillance cameras in these intensifying aggregated crash locations. However, crashes with the hit-and-run in consecutive aggregated and historical aggregated crash locations would decrease the probability of the SI (−0.015 and −0.07, respectively). Meanwhile, hit-and-run crashes in these two types of locations increase the probability of the NI by 0.103 and 0.286, respectively. The heterogeneous results further indicate that the drivers are more likely to run away in severe or minor injury crashes.

Table 8
Estimations and covariances of the intercept terms in hierarchical Bayesian random-intercept logit model with random-effects across groups (spatiotemporal patterns).

Intercept	Spatiotemporal pattern	Severe injury (SI)				Minor Injury (MI)			
		Mean	S.D.	t value	95% BCI	Mean	S.D.	t value	95% BCI
Intercept estimations for each pattern (group-level)									
Pattern1	New aggregated	−3.041	0.222	−13.7	(−3.506, −2.636)	−0.8	0.124	−6.45	(−1.037, −0.561)
Pattern2	Consecutive aggregated	−3.639	0.162	−22.46	(−3.943, −3.32)	−1.082	0.075	−14.43	(−1.232, −0.941)
Pattern3	Intensifying aggregated	−3.376	0.154	−21.92	(−3.658, −3.064)	−0.892	0.068	−13.12	(−1.022, −0.762)
Pattern4	Persistent aggregated	−3.004	0.432	−6.95	(−3.85, −2.172)	−0.665	0.282	−2.36	(−1.214, −0.114)
Pattern5	Sporadic aggregated	−3.254	0.159	−20.47	(−3.547, −2.925)	−0.892	0.076	−11.74	(−1.046, −0.749)
Pattern6	Historical aggregated	−3.063	0.176	−17.4	(−3.386, −2.715)	−0.775	0.09	−8.61	(−0.955, −0.608)
Pattern7	No trend detected	−3.199	0.14	−22.85	(−3.458, −2.912)	−0.892	0.058	−15.38	(−1.005, −0.782)
Pattern8	New dispersed	−2.535	0.546	−4.64	(−3.569, −1.427)	−0.676	0.47	−1.44	(−1.574, 0.235)
Pattern9	Consecutive dispersed	−3.686	0.526	−7.01	(−4.762, −2.66)	−1.062	0.357	−2.97	(−1.712, −0.312)
Pattern10	Sporadic dispersed	−3.039	0.276	−11.01	(−3.587, −2.493)	−0.999	0.18	−5.55	(−1.34, −0.639)
Random effects covariances of the intercepts									
RECov (int1, int1)	Variance for intercept (SI, SI)	0.615	0.311	1.98	(0.191, 1.166)				
RECov (int2, int1)	Covariance between intercepts (SI, MI)	0.052	0.183	0.28	(−0.297, 0.432)				
RECov (int2, int2)	Variance for intercept (MI, MI)	0.511	0.262	1.95	(0.19, 0.989)				

Table 9
Results of the hierarchical Bayesian random-effects-only logit model for pedestrian-injury crashes with the new aggregated pattern.

Variable	Description	Severe injury				Minor injury			
		Mean	S.D.	t value	95% BCI	Mean	S.D.	t value	95% BCI
REMean Int.	Intercept	−39.467	0.94	−41.99	(−41.423, −37.713)	−4.992	0.718	−6.95	(−6.427, −3.628)
REMean PAge2	24 < PedAge ≤ 54 (PedAge ≤ 24)					−4.853	0.488	−9.94	(−5.81, −3.891)
REMean PAge4	PedAge ≥ 65	17.274	0.638	27.08	(15.976, 18.488)				
REMean PAlc2	PedAlcFlag = 'Yes' (No)	3.501	1.103	3.17	(1.367, 5.733)	2.199	0.723	3.04	(0.752, 3.614)
REMean PAlc3	Unknow	9.185	0.726	12.65	(7.764, 10.647)				
REMean Amb2	No ambulance (Yes)	16.032	0.529	30.31	(15.004, 17.125)	−2.89	1.238	−2.33	(−5.182, −0.408)
REMean Grp2	Crossing roadway with vehicle not turning (Walking along roadway)	1.119	0.846	1.32	(−0.477, 2.774)	9.859	0.937	10.52	(8.027, 11.635)
REMean Grp5	Pedestrian in roadway	2.678	0.431	6.21	(1.806, 3.485)				
REMean Grad3	Hillcrest (Level)					29.017	0.744	39	(27.558, 30.525)
REMean Ligh4	Dark – roadway not lighted (Daylight)	20.78	0.643	32.32	(19.527, 22.059)				

Note: The variable in the parentheses is the reference category. REMean denotes posterior mean of the random-effects parameter. Observation number: 432; Burn-In Size: 100,000; Simulation Size: 100,000; Thinning: 20; DIC: 331.21; Hit Probability: 0.968.

In comparison with crashes when pedestrians are walking along the roadways, pedestrians crossing the roadways with a straight vehicle could increase the probability of the SI in all crashes except the crash with new aggregated pattern. For example, a 0.064 increase in the probability of the SI is found in intensifying aggregated crashes. When the pedestrian is crossing the roadways with a turning vehicle, all marginal effects show a decrease in the probability of the SI. For example, consecutive aggregated crashes decrease the probability of the SI by 0.029. Referring to the vehicle maneuver, a plausible reason for the difference effects between vehicles that are going straight and turning might be a lower speed of the turning vehicle. The off-roadway situations show heterogeneous effects on the SI in sporadic aggregated crashes (−0.007) and historical aggregated crashes (0.031). One possible reason for the insignificance of the off-roadways in the whole dataset model might be that the opposite effects in sporadic aggregated and historical aggregated crashes that compensate for each other. When pedestrians are in roadways, all results show an increase in the probability of the SI (0.071 and 0.057 in consecutive and intensifying aggregated crashes, respectively). A similar result could be found in (Mohamed et al., 2013). Dash/dart-out situations increase the risk of the MI in all cases. While heterogeneous results are showed in the SI. An increase of the SI could be found in models with the whole dataset (0.066), the intensifying aggregated dataset (0.019), and the sporadic aggregated dataset (0.012). Meanwhile, a 0.02 decrease in the probability of the SI could be seen in consecutive aggregated crashes. Crashes with a backing vehicle also show heterogeneous results in the SI. The model with the whole dataset decreases the probability of the SI, while crashes with consecutive aggregated and historical aggregated patterns show a slight increase in the probability of the SI. In multiple treat/trapped situations, pedestrians would decrease the risk of SI

Table 10
Results of the hierarchical Bayesian random-effects-only logit model for pedestrian-injury crashes with the consecutive aggregated pattern.

Variable	Description	Severe injury				Minor injury			
		Mean	S.D.	t value	95% BCI	Mean	S.D.	t value	95% BCI
REMean Int.	Intercept	−14.653	0.508	−28.84	(−15.666, −13.685)	−7.718	0.586	−13.17	(−8.861, −6.552)
REMean PAge4	PedAge ≥ 65 (PedAge ≤ 24)	4.793	0.459	10.44	(3.749, 5.574)	2.19	0.415	5.28	(1.307, 2.951)
REMean PAlc2	PedAlcFlag = 'Yes' (No)	4.856	0.253	19.19	(4.367, 5.354)	2.278	0.576	3.95	(1.117, 3.377)
REMean PAlc3	Unknow	12.323	0.443	27.82	(11.48, 13.185)				
REMean PSex2	Female (Male)	−2.963	0.44	−6.73	(−3.827, −2.118)				
REMean Amb2	No ambulance (Yes)	3.401	0.5	6.8	(2.485, 4.402)	7.214	0.473	15.25	(6.22, 8.051)
REMean Run2	Hit and run (No)	−2.912	0.368	−7.91	(−3.614, −2.229)	−2.123	0.459	−4.63	(−2.995, −1.22)
REMean Grp3	Crossing roadway with vehicle turning (Walking along roadway)	−5.909	0.331	−17.85	(−6.581, −5.323)	−2.376	0.49	−4.85	(−3.345, −1.472)
REMean Grp5	Pedestrian in roadway	5.777	0.41	14.09	(5.046, 6.584)				
REMean Grp6	Dash/dart-out	1.531	0.409	3.74	(0.677, 2.282)	6.813	0.538	12.66	(5.835, 7.932)
REMean Grp7	Backing vehicle					−1.477	0.518	−2.85	(−2.473, −0.511)
REMean Grp10	Other/unusual condition					1.305	0.437	2.99	(0.535, 2.203)
REMean Curv2	Curve (Straight)	3.092	0.457	6.77	(2.238, 3.961)				
REMean Grad2	Grade (Level)	1.749	0.543	3.22	(0.676, 2.786)	1.071	0.375	2.86	(0.351, 1.799)
REMean Grad3	Hillcrest	3.629	0.45	8.06	(2.733, 4.506)				
REMean Clas2	Interstate (US route)					−9.475	0.442	−21.44	(−10.335, −8.643)
REMean Clas5	Local street, driveway	−4.117	0.424	−9.71	(−4.828, −3.131)	−2.623	0.457	−5.74	(−3.542, −1.755)
REMean Clas6	Public vehicular area	−13.897	0.415	−33.49	(−14.683, −13.102)	−6.784	0.409	−16.59	(−7.584, −6.022)
REMean Conf3	Two-way, divided (One-way, not divided)	2.52	0.362	6.96	(1.803, 3.194)	1.034	0.455	2.27	(0.096, 1.859)
REMean Ligh3	Dark – lighted roadway (Daylight)	4.814	0.424	11.35	(3.978, 5.594)				
REMean Ligh4	Dark – roadway not lighted	4.254	0.436	9.76	(3.364, 5.036)				
REMean Weat2	Cloudy (Clear)	3.27	0.525	6.23	(2.235, 4.327)				
REMean Hour4	22:00–5:59 (10:00–12:59)					3.132	0.422	7.42	(2.334, 3.966)
REMean Hour5	6:00–9:59	4.283	0.344	12.45	(3.629, 4.961)	1.363	0.347	3.93	(0.655, 1.948)
REMean Cntrl2	Signs (No control)					−1.907	0.489	−3.9	(−2.811, −0.952)

Note: The variable in the parentheses is the base category. REMean denotes posterior mean of the random-effects parameter. Observation number: 3051; Burn-In Size: 80,000; Simulation Size: 20,000; Thinning: 10; DIC: 1878.53; Hit Probability: 0.989.

but increase the risk of MI in models with the whole dataset and sporadic aggregated dataset. In crashes related to taking on/off the bus, the whole dataset model shows a 0.07 increase in the probability of the SI for pedestrians. A crosswalk marking/sign could be set at the end of the bus stop to encourage pedestrians to cross the road behind the bus.

7.3.4. Locality characteristics

In comparison with rural areas, crashes occurred in urban areas decrease the probability of the SI in models with the whole dataset and sporadic/historical aggregated datasets. One possible reason for this is the lower speed limits in urban areas compared to rural areas (Sasidharan et al., 2015; Ulak et al., 2017). Compared to the residential areas, commercial areas show a slight increase of the SI in models with the whole dataset and intensifying/sporadic/historical aggregated datasets. Also, heterogeneous results of the SI could be seen in the lands of farms, woods, and pastures. The models with the whole dataset and sporadic aggregated crash data increase the probability of the SI by 0.019 and 0.1, respectively; while the model for intensifying aggregated crashes decreases the probability of the SI by 0.042.

7.3.5. Roadway characteristics

Compared to straight roadways, curve roadways increase the probability of the SI in models with the whole dataset (0.041) and segmented datasets with consecutive aggregated (0.035) and sporadic aggregated crashes (0.119). In comparison with level roadways, grade roadways increase the probability of the SI in models with the whole dataset (0.036) and segmented datasets with consecutive aggregated (0.011) and intensifying aggregated (0.021) patterns. Heterogeneous results

Table 11
Results of the hierarchical Bayesian random-effects-only logit model for pedestrian-injury crashes with the intensifying aggregated pattern.

Variable	Description	Severe injury				Minor injury			
		Mean	S.D.	t value	95% BCI	Mean	S.D.	t value	95% BCI
REMean Int.	Intercept	−16.47	0.361	−45.62	(−17.123, −15.652)	−5.458	0.345	−15.82	(−6.175, −4.828)
REMean PAge3	55 < PedAge ≤ 64 (PedAge ≤ 24)	0.495	0.334	1.48	(−0.195, 1.06)				
REMean PAge4	PedAge ≥ 65	2.654	0.337	7.88	(2.005, 3.224)				
REMean PAlc2	PedAlcFlag = 'Yes' (No)	5.49	0.449	12.23	(4.561, 6.308)	2.175	0.373	5.83	(1.473, 2.874)
REMean PAlc3	Unknow	7.007	0.398	17.61	(6.215, 7.852)	1.587	0.384	4.13	(0.719, 2.266)
REMean Veh2	Middle (Small)	1.071	0.432	2.48	(0.192, 1.902)				
REMean Veh3	Heavy	2.971	0.337	8.82	(2.346, 3.642)				
REMean Amb2	No ambulance (Yes)	2.201	0.473	4.65	(1.19, 3.118)	4.475	0.297	15.07	(3.841, 4.975)
REMean Run2	Hit and run (No)	2.405	0.335	7.18	(1.813, 3.152)				
REMean DAlc2	Drialcflag = 'yes' (No)	4.938	0.46	10.73	(4.152, 5.76)	2.422	0.318	7.62	(1.786, 3)
REMean DAlc3	Unknow	−3.1	0.395	−7.85	(−3.899, −2.291)				
REMean Grp2	Crossing roadway with vehicle not turning (Walking along roadway)	4.963	0.35	14.18	(4.311, 5.658)				
REMean Grp3	Crossing roadway with vehicle turning	−6.466	0.34	−19.02	(−7.026, −5.744)	−1.64	0.322	−5.09	(−2.29, −1.068)
REMean Grp5	Pedestrian in roadway	3.568	0.361	9.88	(2.943, 4.297)				
REMean Grp6	Dash/dart-out	4.387	0.443	9.9	(3.629, 5.252)	4.167	0.33	12.63	(3.511, 4.868)
REMean Grp10	Other/unusual condition	−0.715	0.424	−1.69	(−1.416, 0.195)	0.741	0.315	2.35	(0.161, 1.42)
REMean Deve2	Commercial (Residential)					−0.501	0.356	−1.41	(−1.145, 0.217)
REMean Deve5	Farms, woods, pastures					8.548	0.378	22.61	(7.833, 9.282)
REMean Grad2	Grade (Level)	1.56	0.397	3.93	(0.759, 2.247)				
REMean Clas2	Interstate (US route)	6.189	0.331	18.7	(5.568, 6.816)				
REMean Clas6	Public vehicular area	−6.086	0.588	−10.35	(−7.083, −5.057)	−3.802	0.308	−12.34	(−4.427, −3.203)
REMean Conf2	Two-way, not divided (One-way, not divided)	−0.826	0.378	−2.19	(−1.584, −0.123)				
REMean Conf3	Two-way, divided	3.756	0.348	10.79	(3.086, 4.342)	0.427	0.298	1.43	(−0.158, 0.991)
REMean Ligh3	Dark – lighted roadway (Daylight)	2.18	0.371	5.88	(1.447, 2.867)	0.968	0.309	3.13	(0.376, 1.531)
REMean Ligh4	Dark – roadway not lighted	3.504	0.407	8.61	(2.73, 4.259)				
REMean Weat2	Cloudy (Clear)					−2.44	0.32	−7.63	(−2.98, −1.819)
REMean Hour2	13:00–16:59 (10:00–12:59)					0.698	0.327	2.13	(0.088, 1.293)
REMean Hour4	22:00–5:59	2.341	0.418	5.6	(1.549, 3.154)				
REMean Cntrl2	Signs (No control)	−2.99	0.468	−6.39	(−3.85, −2.047)	−1.924	0.444	−4.33	(−2.8, −1.156)
REMean Cntrl4	Double yellow line, no passing zone	1.757	0.499	3.52	(0.756, 2.625)				

Note: The variable in the parentheses is the base category. REMean denotes posterior mean of the random-effects parameter. Observation number: 5699; Burn-In Size: 80,000; Simulation Size: 20,000; Thinning: 10; DIC: 3774.66; Hit Probability: 0.991.

of the SI could also be found on roadways at hillcrest and bottom terrains. For roadways at hillcrest terrains, bottom terrains increase the probability of the SI by 0.041 in sporadic aggregated crashes. Previous studies also showed the opposite effects of graded roads on the pedestrian injury severity (Kim et al., 2010; Ma et al., 2018). One possible reason for the heterogeneous result is that curve, grade, crest, and bottom roadways would have bad influence on the vision range and operation of the drivers which would increase the crash injury severity; while drivers are also likely to decrease the speed in these terrains which would decrease the crash injury severity.

In comparison with the U.S. route, all results in interstates show an increase in the probability of the SI. As most interstates are highways, over 50% of the crashes result in SI for pedestrians in the original dataset. It is noted that interstate routes increase the probability of the SI by 0.131 and 0.857 in locations with intensifying aggregated and historical aggregated crashes, respectively. The historical temporal pattern indicates that these locations are no longer crash-prone areas. Hence, more attention should be focused on interstates in locations with intensifying aggregated crashes. Meanwhile, state secondary routes, local street driveways, and public vehicular areas all decrease the probability of the SI for pedestrians. The probability of the SI decreases with the decrease of the roadway hierarchy, and this might be caused by the decrease of the

Table 12
Results of the hierarchical Bayesian random-effects-only logit model for pedestrian-injury crashes with the sporadic aggregated pattern.

Variable	Description	Severe injury				Minor injury			
		Mean	S.D.	t value	95% BCI	Mean	S.D.	t value	95% BCI
REMean Int.	Intercept	−17.356	0.54	−32.14	(−18.401, −16.31)	−9.573	0.45	−21.27	(−10.448, −8.667)
REMean Age2	24 < PedAge ≤ 54 (PedAge ≤ 24)	1.115	0.422	2.64	(0.237, 1.912)				
REMean Age3	55 < PedAge ≤ 64	4.237	0.469	9.03	(3.311, 5.104)				
REMean Age4	PedAge ≥ 65	4.458	0.559	7.97	(3.469, 5.611)				
REMean Alc2	PedAlcFlag = 'Yes' (No)	7.243	0.573	12.64	(6.154, 8.33)	−1.166	0.655	−1.78	(−2.33, 0.166)
REMean Alc3	Unknow	11.264	0.514	21.91	(10.218, 12.291)	−5.498	0.434	−12.67	(−6.406, −4.689)
REMean Veh3	Heavy	3.709	0.595	6.23	(2.648, 4.869)				
REMean Amb2	No ambulance (Yes)	0.796	0.515	1.55	(−0.111, 1.889)	6.051	0.448	13.51	(5.283, 6.992)
REMean DAlc2	Drialcflag = 'yes' (No)	2.979	0.469	6.35	(2.129, 3.916)				
REMean Grp2	Crossing roadway with vehicle not turning (Walking along roadway)	1.403	0.565	2.48	(0.282, 2.52)				
REMean Grp3	Crossing roadway with vehicle turning	−4.827	0.503	−9.6	(−5.982, −3.945)				
REMean Grp4	Off roadway					3.834	0.441	8.69	(3.011, 4.685)
REMean Grp6	Dash/dart-out	3.774	0.495	7.62	(2.807, 4.772)	7.666	0.602	12.73	(6.54, 8.83)
REMean Grp8	Multiple threat/trapped					1.381	0.362	3.81	(0.753, 2.147)
REMean Grp10	Other/unusual condition					1.702	0.396	4.3	(0.932, 2.493)
REMean Urb2	Urban (Rural)	−2.676	0.534	−5.01	(−3.766, −1.574)				
REMean Deve2	Commercial (Residential)	1.814	0.43	4.22	(0.953, 2.574)				
REMean Deve5	Farms, woods, pastures	7.824	0.525	14.9	(6.856, 8.87)	4.392	0.396	11.09	(3.628, 5.151)
REMean Curv2	Curve (Straight)	8.045	0.506	15.9	(7.077, 9.091)				
REMean Grad4	Bottom (Level)	3.726	0.599	6.22	(2.561, 4.863)	2.054	0.583	3.52	(0.985, 3.273)
REMean Clas6	Public vehicular area (US route)	−6.659	0.547	−12.17	(−7.676, −5.674)	−3.749	0.41	−9.14	(−4.629, −3.027)
REMean Conf3	Two-way, divided (One-way, not divided)	−0.5	0.502	−1	(−1.492, 0.521)				
REMean Ligh3	Dark – lighted roadway (Daylight)	4.866	0.391	12.45	(4.101, 5.572)				
REMean Ligh4	Dark – roadway not lighted	4.573	0.488	9.37	(3.643, 5.553)	0.712	0.52	1.37	(−0.574, 1.634)
REMean Hour5	6:00–9:59 (10:00–12:59)	3.047	0.485	6.28	(2.119, 3.996)				
REMean Cntrl2	Signs (No control)	−4.722	0.702	−6.73	(−6.155, −3.542)				

Note: The variable in the parentheses is the base category. REMean denotes posterior mean of the random-effects parameter. Observation number: 2533; Burn-In Size: 80,000; Simulation Size: 20,000; Thinning: 10; DIC: 1552.43; Hit Probability: 0.989.

speed limits in lower-hierarchy roadways (Li and Fan, 2019b). Compared to the one-way road without the central divider, both two-way roads with and without central dividers increase the probability of the SI in the whole dataset model, while heterogeneous results could be seen in models with segmented crash dataset. For two-way roads without central dividers, crashes with intensifying aggregated pattern decrease the probability of the SI by −0.01. For two-way roads with central dividers, consecutive and intensifying aggregated crashes increase the probability of the SI by 0.018 and 0.045, respectively; while sporadic aggregated crashes slightly decrease the probability of the SI by 0.004.

7.3.6. Environment and time characteristics

In comparison with the daylight condition, the dawn/dusk condition could slightly increase the SI (0.001) in the whole dataset model. Both dark with and without roadway light conditions increase the probability of the SI. The dark without roadway light conditions result in severer injuries than dark with roadway light conditions in all models. For dark without roadway light conditions, locations with new, consecutive, and intensifying aggregated crashes increase the probability of the SI by 0.124, 0.046, and 0.055, respectively. The results indicate an urgent need to set roadside light in these locations and these findings are in accordance with (Yasmin et al., 2014). Compared to the clear weather, the cloudy weather increases

Table 13
Results of the hierarchical Bayesian random-effects-only logit model for pedestrian-injury crashes with the historical aggregated pattern.

Variable	Description	Severe injury				Minor injury			
		Mean	S.D.	t value	95% BCI	Mean	S.D.	t value	95% BCI
REMean Int.	Intercept	−1.408	0.558	−2.52	(−2.48, −0.332)	1.524	0.466	3.27	(0.598, 2.418)
REMean Age2	24 < PedAge ≤ 54 (PedAge ≤ 24)	−2.953	0.547	−5.4	(−4.008, −1.88)	−3.392	0.479	−7.08	(−4.3, −2.424)
REMean Age3	55 < PedAge ≤ 64					−5.089	0.645	−7.89	(−6.29, −3.764)
REMean Alc2	PedAlcFlag = 'Yes' (No)	2.508	0.632	3.97	(1.238, 3.669)	8.578	0.516	16.62	(7.466, 9.474)
REMean Alc3	Unknow	14.267	0.647	22.05	(12.931, 15.519)				
REMean Veh3	Heavy (Small)	1.986	0.852	2.33	(0.32, 3.587)				
REMean Amb2	No ambulance (Yes)	5.574	0.605	9.21	(4.444, 6.76)	6.198	0.429	14.45	(5.388, 7.123)
REMean Run2	Hit and run	−14.834	0.605	−24.52	(−15.998, −13.618)	−5.997	0.489	−12.26	(−6.916, −5.025)
REMean Grp2	Crossing roadway with vehicle not turning (Walking along roadway)					−4.072	0.521	−7.82	(−5.137, −3.077)
REMean Grp3	Crossing roadway with vehicle turning	−10.201	0.354	−28.82	(−10.877, −9.492)	−4.953	0.532	−9.31	(−5.993, −3.937)
REMean Grp4	Off roadway					−9.072	0.538	−16.86	(−10.12, −8.009)
REMean Grp7	Backing vehicle					−8.206	0.615	−13.34	(−9.383, −6.998)
REMean Grp10	Other/unusual condition					−6.781	0.548	−12.37	(−7.736, −5.627)
REMean Urb2	Urban (Rural)	−11.314	0.679	−16.66	(−12.536, −9.866)				
REMean Deve2	Commercial (Residential)					−1.679	0.447	−3.76	(−2.65, −0.899)
REMean Clas2	Interstate (US route)	27.588	0.719	38.37	(26.32, 29.096)				
REMean Clas6	Public vehicular area	−11.848	0.469	−25.26	(−12.739, −10.899)	−4.699	0.979	−4.8	(−6.677, −2.865)
REMean Ligh4	Dark – roadway not lighted (Daylight)					−4.587	0.445	−10.31	(−5.416, −3.691)
REMean Hour4	22:00–5:59 (10:00–12:59)	7.763	0.391	19.85	(6.965, 8.493)	2.706	0.388	6.97	(1.953, 3.484)

Note: The variable in the parentheses is the base category. REMean denotes posterior mean of the Random-effects parameter. Observation number: 1121; Burn-In Size: 150,000; Simulation Size: 20,000; Thinning: 10; DIC: 721.11; Hit Probability: 0.988.

the SI by 0.034 and 0.014 in crashes with consecutive aggregated and intensifying aggregated patterns, respectively. A similar result could be referred to (Aziz et al., 2013).

Compared to the "noon" period (10:00–12:59), the "morning" period (6:00–9:59) increases the probability of the SI for pedestrians (up to 0.04 in crashes with consecutive aggregated pattern). For the "afternoon" period (13:00–16:59), crashes with intensifying aggregated pattern decrease the probability of the SI by 0.005 and increase the MI by 0.058. The "early evening" period (17:00–21:59) is found to increases the probability of the SI and MI by 0.006 and 0.029 in the whole dataset model. The "deep night" period (17:00–21:59) increases the probability of the SI in the whole dataset model (0.043), the intensifying aggregated crash model (0.031), and the historical aggregated crash model (0.05). However, locations with consecutive aggregated crashes show a 0.013 probability decrease for the SI and a 0.218 increase for the MI. Hence, roadside lights or electronic traffic signs to alert drivers to watch out for pedestrians are needed for these locations. The variations and heterogeneous results of the injury severities further emphasize the importance to consider the temporal variations.

7.3.7. Traffic control characteristics

In comparison with the no traffic control situation, traffic sign controls in the model with the whole dataset and segmented datasets with intensifying and sporadic aggregated crashes all present a decrease in the probability of the SI. However, a 0.007 and −0.085 probability changes in the SI and MI could be found in in locations with consecutive aggregated crashes. Possible reasons for the heterogeneous results could be: 1) traffic signs could alert the driver and mitigate the crash severity level to some extent; 2) alert signs are always set in crash-prone locations with severe crashes; and 3) locations with consecutive aggregated crashes are more likely to result in the SI for pedestrians. Kim et al., (2010) also observed heterogeneous effects of the traffic sign control on pedestrian injury severities. In the whole dataset model, both signal and human control situations decrease the probability of the SI by −0.017 and −0.073, respectively. For double yellow line or no passing zone situations, the whole dataset model and the intensifying aggregated model increase the probability of the SI by 0.006 and 0.025, respectively. These further indicate the hazard of crossing these roadways, and the need of a safer control scheme, such as a stop line with the crosswalk to guide the pedestrian to cross these locations safely.

Table 14
Marginal effects of significant variables in models with the whole dataset and spatiotemporal segmented dataset.

Dataset	Whole data			New aggregated			Consecutive aggregated			Intensifying aggregated			Sporadic aggregated			Historical aggregated		
Severity level	SI	MI	NI	SI	MI	NI	SI	MI	NI	SI	MI	NI	SI	MI	NI	SI	MI	NI
24 < PedAge ≤ 54 (PedAge ≤ 24)	0.007	−0.027	0.02	−0.028	0.003	0.025							0.009	−0.001	−0.008	−0.008	−0.16	0.168
55 < PedAge ≤ 64	0.057	−0.032	−0.025							0.006	−0.004	−0.003	0.051	−0.005	−0.046	0.021	−0.217	0.197
PedAge ≥ 65	0.095	0.003	−0.098	0.091	−0.021	−0.07	0.04	0.108	−0.148	0.039	−0.022	−0.017	0.058	−0.006	−0.052			
PedAlcFlag = 'Yes' (No)	0.061	0.031	−0.091	0.031	0.027	−0.058	0.036	0.125	−0.161	0.067	0.155	−0.222	0.059	−0.036	−0.023	−0.031	0.419	−0.388
PedSex = Female (Male)	−0.008	−0.009	0.017				−0.021	0.009	0.012									
Middle (Small)	0.025	−0.014	−0.011							0.013	−0.008	−0.006						
Heavy	0.143	−0.023	−0.12							0.047	−0.026	−0.021	0.044	−0.005	−0.039	0.017	−0.01	−0.007
No ambulance (Yes)	0.098	0.154	−0.252	0.028	−0.03	0.002	0.012	0.211	−0.223	0.005	0.244	−0.248	0.003	0.123	−0.126	0.028	0.25	−0.278
Hit and run (No)	0.007	−0.026	0.019				−0.015	−0.088	0.103	0.035	−0.019	−0.016				−0.07	−0.215	0.286
DriAlcFlag = 'Yes' (No)	0.076	0.003	−0.079							0.057	0.174	−0.23	0.033	−0.004	−0.029			
Crossing roadway with vehicle not turning (Walking along roadway)	0.086	−0.004	−0.083	−0.012	0.884	−0.872				0.064	−0.042	−0.022	0.012	0.007	−0.019	0.018	−0.175	0.156
Crossing roadway with vehicle turning	−0.07	−0.006	0.075				−0.029	−0.098	0.126	−0.038	−0.102	0.14	−0.029	0.004	0.025	−0.053	−0.19	0.243
Off Roadway													−0.007	0.293	−0.286	0.031	−0.339	0.308
Pedestrian in roadway	0.076	−0.042	−0.034	0.025	−0.002	−0.022	0.071	−0.033	−0.038	0.057	−0.031	−0.027						
Dash/dart-out	0.066	0.082	−0.148				−0.02	0.516	−0.497	0.019	0.392	−0.411	0.012	0.663	−0.675			
Backing vehicle	−0.027	−0.033	0.059				0.005	−0.067	0.062							0.027	−0.305	0.278
Multiple threat/trapped	−0.02	0.074	−0.054										−0.002	0.064	−0.062			
Bus related vehicle	0.07	0.024	−0.094															
Urban (Rural)	−0.036	−0.026	0.062										−0.028	0.003	0.025	−0.261	0.129	0.132
Commercial (Residential)	0.007	−0.028	0.02							0.004	−0.041	0.038	0.015	−0.002	−0.013	0.008	−0.079	0.071
Farms, woods, pastures	0.019	−0.011	−0.008							−0.042	0.682	−0.64	0.1	0.354	−0.454			
Curve (Straight)	0.041	−0.023	−0.018				0.035	−0.017	−0.018				0.119	−0.013	−0.106			
Grade (Level)	0.036	−0.001	−0.035				0.011	0.054	−0.064	0.021	−0.012	−0.009						
Hillcrest				−0.024	0.84	−0.816	0.041	−0.02	−0.021									
Bottom	−0.027	0.1	−0.073										0.041	0.111	−0.152			
Interstate (US route)	0.077	−0.043	−0.034				0.019	−0.2	0.18	0.131	−0.066	−0.065				0.857	−0.342	−0.515
State secondary route	−0.026	0.015	0.011															
Local street, driveway	−0.047	0.026	0.02				−0.026	−0.149	0.175									
Public vehicular area	−0.098	−0.024	0.122				−0.046	−0.262	0.309	−0.034	−0.205	0.239	−0.035	−0.109	0.144	−0.071	−0.179	0.25
Two-way, not divided (One-way, not divided)	0.026	−0.015	−0.011							−0.01	0.006	0.004						
Two-way, divided	0.059	0.005	−0.064				0.018	0.048	−0.066	0.045	0.006	−0.051	−0.004	−0.008	0.012			
Dawn/dusk light (Daylight)	0.01	−0.038	0.028															
Dark – lighted roadway	0.038	−0.021	−0.017				0.046	−0.021	−0.025	0.02	0.068	−0.087	0.045	−0.004	−0.04			
Dark – roadway not lighted	0.067	−0.038	−0.03	0.124	−0.019	−0.105	0.046	−0.019	−0.027	0.055	−0.03	−0.025	0.05	0.019	−0.069	0.021	−0.186	0.165
Cloudy (Clear)							0.034	−0.016	−0.017	0.014	−0.151	0.136						

(*continued on next page*)

Table 14 (*continued*)

Dataset	Whole data			New aggregated			Consecutive aggregated			Intensifying aggregated			Sporadic aggregated			Historical aggregated		
Severity level	SI	MI	NI	SI	MI	NI	SI	MI	NI	SI	MI	NI	SI	MI	NI	SI	MI	NI
Snow, sleet, hail, freezing rain/drizzle	−0.089	0.052	0.037															
13:00–16:59 (10:00–12:59)										−0.005	0.058	−0.053						
17:00–21:59	0.006	0.029	−0.035															
22:00–5:59	0.043	0.047	−0.09				−0.013	0.218	−0.205	0.031	−0.017	−0.014				0.054	0.094	−0.149
6:00–9:59	0.026	0.013	−0.039				0.04	0.054	−0.093				0.034	−0.004	−0.03			
Signs (No control)	−0.034	−0.029	0.063				0.007	−0.085	0.078	−0.019	−0.113	0.132	−0.029	0.004	0.025			
Signal	−0.017	0.01	0.007															
Double yellow line, no Passing zone	0.006	−0.022	0.016							0.025	−0.014	−0.011						
Human control	−0.073	−0.029	0.102															

Note: Variable in the parentheses is the base category. Severe injury (SI), minor injury (MI), and no injury (NI).

8. Conclusions

Using data of pedestrian-vehicle crashes involved single vehicle in North Carolina from 2007 to 2018, this study explores factors affecting pedestrian-injury severities considering the spatiotemporal patterns of the crashes. Ten spatiotemporal patterns of the crash are identified by conducting an improved spatiotemporal analysis. Significant temporal instability and the spatiotemporal instability of the factors to the pedestrian-injury crashes are identified by the likelihood ratio tests. Then the performances of different hierarchical Bayesian random-effects logit (HBREL) models are compared. A hierarchical Bayesian random intercept logit model with random-effects across groups (HBRIL-grp) and a hierarchical Bayesian random-effects-only logit (HBREOL) model with random-effects across observations are utilized to investigate significant contributing factors of pedestrian-injury severities such as human, vehicle, crash, locality, roadway, environment, time, and traffic control characteristics for the whole dataset and each segmented dataset, respectively. Marginal effects are further calculated for better interpreting the impacts of categorical variables on injury severities considering different spatiotemporal features of the crashes.

To account for heterogeneity across the spatiotemporal pattern groups within one model simultaneously, a HBRIL-grp model is employed with the whole dataset. Results of model comparisons indicate that accounting for random-effects across observations and increasing the number of random parameters could both improve the performance of the model fit and prediction accuracy. Hence, to further account for the unobserved heterogeneity across observations in spatiotemporal segmented datasets, the commonly used HBRIL is modified into a HBREOL model by allowing all parameters to be randomly distributed across observations. Compared to the traditional Bayesian multinomial logit model and random intercept model (hit probability is around 50% and 70%, respectively), the significant improvements in the DIC and the 99% hit probability again prove the superiority of the HBREOL models.

The findings of this study underscore the importance of accounting for the spatiotemporal patterns of the dataset and the random-effects across observations/groups. Heterogeneous marginal effect results are identified between different spatiotemporal crashes. These findings also provide insights into the inherent reason for the temporal instability/tendency of the crash and correlated factors (Behnood and Mannering, 2016). More attention should be given to locations in which the spatially aggregated crashes have new, consecutive, and intensifying temporal tendencies. For example, factors such as alcohol involvement, dark without roadside light, pedestrians on the roadways, and interstates all significantly increase the probability of the SI in locations with new, consecutive, and intensifying aggregated crashes. More restricted regulations and severe penalties are needed for alcohol-involved and hit-and-run crashes. Also, more frequent patrols and alcohol tests are needed for these locations with intensifying aggregated crashes. What's more, it is suggested to lower speed limits and restrict the permitted traveling time for trucks. The permitted traveling time for the trucks is supposed to be different from the traveling peak hours of the pedestrians.

This paper provides a framework for engineers and researchers to explore the factors to crashes with different spatiotemporal patterns, and gives specific countermeasures in crash-prone areas. The random-effects-only model could well help to identify the contributing factors for each crash observation. However, considering the limited size and factors in the existing crash data, a dynamically updated model with a less time-consuming estimation process would be preferred in giving a real-time safety alert to drivers/pedestrians.

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 authors want to express their deepest gratitude for the financial support from the United States Department of Transportation, University Transportation Center through the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE) at the University of North Carolina at Charlotte (Grant Number: 69A3551747133).

References

Abay, K.A., 2013. Examining pedestrian-injury severity using alternative disaggregate models. Research in Transportation Economics 43, 123–136.
Aziz, H.M.A., Ukkusuri, S.V., Hasan, S., 2013. Exploring the determinants of pedestrian-vehicle crash severity in New York City. Accident Analysis and Prevention 50, 1298–1309.
Beck, L.F., Dellinger, A.M., O'Neil, M.E., 2007. Motor vehicle crash injury rates by mode of travel, United States: Using exposure-based methods to quantify differences. American Journal of Epidemiology 166, 212–218.
Behnood, A., Mannering, F., 2017. Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research 16, 35–47.
Behnood, A., Mannering, F., 2016. An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. Analytic Methods in Accident Research 12, 1–17.
Bhat, C.R., Astroza, S., Lavieri, P.S., 2017. A new spatial and flexible multivariate random-coefficients model for the analysis of pedestrian injury counts by severity level. Analytic Methods in Accident Research 16, 1–22.
Blazquez, C.A., Picarte, B., Calderón, J.F., Losada, F., 2018. Spatial autocorrelation analysis of cargo trucks on highway crashes in Chile. Accident Analysis and Prevention 120, 195–210.

Chen, C., Zhang, G., Liu, X.C., Ci, Y., Huang, H., Ma, J., Chen, Y., Guan, H., 2016. Driver injury severity outcome analysis in rural interstate highway crashes: a two-level Bayesian logistic regression interpretation. Accident Analysis and Prevention 97, 69–78.
Chen, C., Zhang, G., Tian, Z., Bogus, S.M., Yang, Y., 2015. Hierarchical Bayesian random intercept model-based cross-level interaction decomposition for truck driver injury severity investigations. Accident Analysis and Prevention 85, 186–198.
Chen, Z., Fan, W. (David), 2019. A multinomial logit model of pedestrian-vehicle crash severity in North Carolina. International Journal of Transportation Science and Technology 8, 43–52.
Dai, D., 2012. Identifying clusters and risk factors of injuries in pedestrian-vehicle crashes in a GIS environment. Journal of Transport Geography 24, 206–214.
Ebdon, D., 1985. Statistics in Geography. Oxford Oxfordshire, New York.
Getis, A., Ord, J.K., 2010. The Analysis of Spatial Association by Use of Distance Statistics. Perspectives on Spatial Data Analysis. Springer, Berlin, Heidelberg, pp. 127–145.
Greene, W.H., Hensher, D.A., Rose, J.M., 2005. Using Classical Simulation-Based Estimators to Estimate Individual WTP Values, Applications of Simulation Methods in Environmental and Resource Economics. Springer, Dordrecht, pp. 17–33.
Gudes, O., Varhol, R., Sun, Q., (Chayn), Meuleners, L.,, 2017. Investigating articulated heavy-vehicle crashes in western Australia using a spatial approach. Accident Analysis and Prevention 106, 243–253.
Haleem, K., Alluri, P., Gan, A., 2015. Analyzing pedestrian crash injury severity at signalized and non-signalized locations. Accident Analysis and Prevention 81, 14–23.
Hu, L., Wu, X., Huang, J., Peng, Y., Liu, W., 2020. Investigation of clusters and injuries in pedestrian crashes using GIS in Changsha, China. Safety Science 127, 104710.
Huang, H., Chin, H.C., Haque, M.M., 2008. Severity of driver injury and vehicle damage in traffic crashes at intersections: A Bayesian hierarchical analysis. Accident Analysis and Prevention 40 (1), 45–54.
Kendall, M.G., Gibbons, J.D., 1990. Rank correlation methods. Griffin, London.
Kim, J., Ulfarsson, G.F., Shankar, V.N., Mannering, F., 2010. A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. Accident Analysis and Prevention 42 (6), 1751–1758.
Li, D., Ranjitkar, P., Zhao, Y., Yi, H., Rashidi, S., 2017. Analyzing pedestrian crash injury severity under different weather conditions. Traffic Injury Prevention 18 (4), 427–430.
Li, Y., Abdel-Aty, M., Yuan, J., Cheng, Z., Lu, J., 2020. Analyzing traffic violation behavior at urban intersections: A spatio-temporal kernel density estimation approach using automated enforcement system data. Accident Analysis and Prevention 141, 105509.
Li, Y., Fan, W. (David), 2019a. Modelling severity of pedestrian-injury in pedestrian-vehicle crashes with latent class clustering and partial proportional odds model: A case study of North Carolina. Accident Analysis and Prevention 131, 284–296.
Li, Y., Fan, W. (David), 2019b. Pedestrian injury severities in pedestrian-vehicle crashes and the partial proportional odds logit model: Accounting for age difference. Transportation Research Record 2673 (5), 731–746.
Li, Z., Chen, C., Ci, Y., Zhang, G., Wu, Q., Liu, C., Qian, Z., (Sean),, 2018. Examining driver injury severity in intersection-related crashes using cluster analysis and hierarchical Bayesian models. Accident Analysis and Prevention 120, 139–151.
Li, Z., Chen, X., Ci, Y., Chen, C., Zhang, G., 2019. A hierarchical Bayesian spatiotemporal random parameters approach for alcohol/drug impaired-driving crash frequency analysis. Analytic Methods in Accident Research 21, 44–61.
Liu, J., Hainen, A., Li, X., Nie, Q., Nambisan, S., 2019. Pedestrian injury severity in motor vehicle crashes: An integrated spatio-temporal modeling approach. Accident Analysis and Prevention 132, 105272.
Ma, Z., Lu, X., Chien, S.I.J., Hu, D., 2018. Investigating factors influencing pedestrian injury severity at intersections. Traffic Injury Prevention 19 (2), 159–164.
Mann, H.B., 1945. Nonparametric tests against trend. Econometrica 13, 245.
Mannering, F., 2018. Temporal instability and the analysis of highway accident data. Analytic Methods in Accident Research 17, 1–13.
Mannering, F.L., Bhat, C.R., 2014. Analytic methods in accident research: Methodological frontier and future directions. Analytic Methods in Accident Research 1, 1–22.
Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic Methods in Accident Research 11, 1–16.
Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. Accident Analysis and Prevention 39 (3), 459–468.
Mohamed, M.G., Saunier, N., Miranda-Moreno, L.F., Ukkusuri, S.V., 2013. A clustering regression approach: A comprehensive injury severity analysis of pedestrian-vehicle crashes in New York, US and Montreal, Canada. Safety Science 54, 27–37.
Mokhtarimousavi, S., 2019. A time of day analysis of pedestrian-involved crashes in California: investigation of injury severity, a logistic regression and machine learning approach using HSIS data. Institute of Transportation Engineers. ITE Journal 89 (10), 25–33.
Mokhtarimousavi, S., Anderson, J.C., Azizinamini, A., Hadi, M., 2020. Factors affecting injury severity in vehicle-pedestrian crashes: A day-of-week analysis using random parameter ordered response models and Artificial Neural Networks. International Journal of Transportation Science and Technology 9 (2), 100–115.
Moran, P.A.P., 1948. The interpretation of statistical maps. Journal of the Royal Statistical Society: Series B 10 (2), 243–251.
Mujalli, R.O., Garach, L., López, G., Al-Rousan, T., 2019. Evaluation of injury severity for pedestrian-vehicle crashes in Jordan using extracted rules. Journal of Transportation Engineering Part A 145 (7), 04019028.
Munira, S., Sener, I.N., Dai, B., 2020. A Bayesian spatial Poisson-lognormal model to examine pedestrian crash severity at signalized intersections. Accident Analysis and Prevention 144, 105679.
NHTSA, 2019. Traffic Safety Facts 2017 Data: Pedestrians. USDOT, Washington, DC, pp. 1–11.
Ouni, F., Belloumi, M., 2018. Spatio-temporal pattern of vulnerable road user's collisions hot spots and related risk factors for injury severity in Tunisia. Transportation Research Part F 56, 477–495.
Plug, C., Xia, J., Caulfield, C., 2011. Spatial and temporal visualisation techniques for crash analysis. Accident Analysis and Prevention 43 (6), 1937–1946.
Rossi, P.E., Allenby, G.M., McCulloch, R., 2012. Bayesian Statistics and Marketing. John Wiley & Sons.
Sasidharan, L., Wu, K.F., Menendez, M., 2015. Exploring the application of latent class cluster analysis for investigating pedestrian crash injury severities in Switzerland. Accident Analysis and Prevention 85, 219–228.
Song, L., Fan, W., 2020. Combined latent class and partial proportional odds model approach to exploring the heterogeneities in truck-involved severities at cross and T-intersections. Accident Analysis and Prevention 144, 105638.
Songchitruksa, P., Zeng, X., 2010. Getis-ord spatial statistics to identify hot spots by using incident management data. Transportation Research Record 2165 (1), 42–51.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B 64 (4), 583–639.
Sun, M., Sun, X., Shan, D., 2019. Pedestrian crash analysis with latent class clustering method. Accident Analysis and Prevention 124, 50–57.
Ulak, M.B., Kocatepe, A., Ozguven, E.E., Horner, M.W., Spainhour, L., 2017. Geographic information system–based spatial and statistical analysis of severe crash hotspot accessibility to hospitals. Transportation Research Record 2635 (1), 90–97.
Washington, S., Karlaftis, M., Mannering, F., 2011. Statistical and Econometric Methods for Transportation Data Analysis. Chapman and Hall/CRC, Boca Raton, FL.
Yalcin, G., Duzgun, H.S., 2015. Spatial analysis of two-wheeled vehicles traffic crashes: Osmaniye in Turkey. KSCE Journal of Civil Engineering 19 (7), 2225–2232.
Yasmin, S., Eluru, N., Ukkusuri, S.V., 2014. Alternative ordered response frameworks for examining pedestrian injury severity in New York City. Journal of Transportation Safety and Security 6 (4), 275–300.