Research Article
Safety Analysis of Signal-Coordinated Arterials with Multiple Heterogeneities

Yingfei Fan 1,2,3, Guopeng Zhang 4, Zhixuan Jia 1, and Minjie Jin 1

1 School of Transportation and Logistics Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China
2 National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Chengdu 611756, China
3 Department of Civil and Environmental Engineering, National University of Singapore, Singapore 117576
4 College of Engineering, Zhejiang Normal University, Jinhua 321004, China

Correspondence should be addressed to Yingfei Fan; fanyf@my.swjtu.edu.cn

Received 26 April 2021; Revised 18 July 2021; Accepted 17 August 2021; Published 10 September 2021

1. Introduction

Arterials obtain growing concerns with the increase of the major traffic flow among urban road networks. For ensuring operational efficiency, the control technique of signal coordination was widely applied along the arterials in recent decades [1–3]. By coupling the intersections simultaneously along arterials, signal coordination generates the “green wave” by optimizing the signal attributes of phase, timing, and cycle length [4]. The advantages of coordinated signalization are verified from several aspects, such as increasing the traffic output, alleviating the vehicular delay, and decreasing the phenomenon of “stop and go” [3].

The previous study focuses mainly on the efficiency improvement of arterials with coordinated signalization. It should be noted that the efficiency and safety impacts are not consistent. Signal coordination may induce safety hazards along arterials with the efficiency optimization. Since safety is the precondition of sustainable operation, it is imperative to provide an in-depth safety exploration of coordinated arterials with appropriate models.

The safety evaluation model is a useful tool to identify the crash risks [5, 6]. During the modeling procedure, the heterogeneity should be nonnegligible on account of the connected traffic flow along signal-coordinated arterials [7]. Heterogeneous correlations, posing an important effect on safety analysis results, stand for the interference involved in the data source of coordinated arterials. The micro-aspect (i.e., on the roadway segments/or at the signalized intersections), there is the multivariate heterogeneity between injury and property damage only (PDO) crashes, since crashes from different severities may show the diverse contributors [8]. Meanwhile, signal coordination results in a strong association among vehicles with the “green wave,” thus forming the structural heterogeneity between crashes in each traffic analysis zone (TAZ) at the meso-viewpoint [9]. Regarding the macrolevel (i.e., along the whole coordinated

In the corresponding research available, the safety impact remains controversial in implementing signal coordination on arterials, which calls for an in-depth exploration with the appropriate statistical methods. Based on the traffic data from Ann Arbor City (Michigan, USA), the paper proposes a safety evaluation model considering the multiple heterogeneities. In terms of arterials with the coordinated signalization, modeling results imply that (1) the multivariate heterogeneity shows the strongest interaction on crash frequency, followed by the spatiotemporal and structural heterogeneities, and (2) the spatial variation is unrelated to the temporal change among crashes in the denoted traffic analysis zones (TAZs). In an attempt to alleviate the coupled crash risks along the coordinated arterials, the study emphasizes the necessity of dividing the subcontrol traffic areas in real time according to the correlated degree of crash distribution. Meanwhile, the modeling framework with multiple heterogeneities can be applied for the safety analysis of other urban roads.
arterials), it calls for the need to distinguish the specific form of spatiotemporal heterogeneity among the TAZs [10].

The innovation of the paper is to provide the comprehensive safety analysis of signal-coordinated arterials with multiple heterogeneities. The paper begins with the literature review in Section 2. Section 3 divides the signal-coordinated arterials into roadway segments, signalized intersections, and TAZs. Each corresponding area retrieves the data of road design features, traffic characteristics, and crash counts. Section 4 describes seven statistical models, including Poisson log-normal (PLN) model, multivariate Poisson log-normal (MPLN) model, joint multivariate Poisson log-normal (JMPLN) model, joint multivariate Poisson log-normal conditional autoregressive (JMPLN-CAR) model, JMPLN-CAR model with fixed spatial and linear temporal variation, JMPLN-CAR model with fixed spatial and linear-quadratic temporal variation, and JMPLN-CAR model with time-varying spatial interaction. To compare the predictable performance of models, the paper adopts three kinds of criteria, i.e., classical criteria (mean absolute deviance (MAD), the weighted mean absolute deviance (WMAD), mean squared prediction error (MSPE), and weighted mean squared prediction error (WMSPE)), prevalent criteria (deviance information criterion (DIC) and total deviance information criterion (TDIC)), and emerging criteria (log pseudo marginal likelihoods (LPML) and total log pseudo marginal likelihoods (TLPML)). In addition, variation proportion is feasible to reveal the modeling fitness explained by the heterogeneous correlations. With the evaluated results (Section 5), safety improvement suggestions are outlined in Section 6.

2. Literature Review

2.1. Efficiency Optimization of Signal-Coordinated Arterials. The research of signal coordination draws surging attention from scholars. In contrast to previous control modes (e.g., the fixed period and isolated signal), the coordinated signalization systematically integrates arterials as a whole with advanced approaches (e.g., the coevolutionary analysis, sensor technology, automatic control, and information theory) [11], thus significantly improving the movement efficiency of vehicles.

Specifically, the efficiency optimization of signal coordination had two aspects: (1) setting the green bandwidth to increase the traffic output [12] and (2) quantifying the performance index, such as reducing the vehicular delay, decreasing the queue length, considering the product of traffic flow, and improving the traveling speed [13]. To achieve the mentioned optimization objects, several algorithms were provided of the Max-Band control [14], forward-backward [15, 16], fuzzy control, game theory [17], ant colony [18], hierarchical and distributed control [19], and linear programming of mixed integers [14].

In summary, the previous literature focuses mainly on the efficiency improvement of signal coordination. It needs
to provide the corresponding safety evaluation since the efficiency and safety impacts are not consistent.

### 2.2. Safety Evaluation of Signal-Coordinated Arterials.

Signal coordination is able to optimize the moving efficiency of vehicles, while its safety impact is controversial in the current studies. One viewpoint is that the safety status improves significantly after applying the signal coordination on arterials. Carter and St-Onge [20] verified the 4.9% crash reduction of Phoenix city (Arizona, USA) under the operation of signal-coordinated arterials. Meanwhile, the unsafe behaviors were decreased in the study of Khattak et al. [21] due to the existence of consecutively vehicular platoons with signal coordination. Moreover, scholars identified the decrease of pedestrian, right-turn, and right-angle crashes with the control of signal coordination in the before-after study [22, 23]. An explanation was that signal coordination mitigates the phenomenon of “stop and go,” therefore decreasing the possibility of crash occurrence [24, 25].

In contrary to the above findings, there is a conclusion that signal coordination incurs safety hazards along arterials. For instance, crashes occurring at the signalized intersections were migrated to the roadway segments due to the performance of signal coordination [9]. To alleviate the vehicular delay, signal coordination was also allocated the high-speed limit on arterials. However, this countermeasure was potentially correlated with severe injury outcomes [26]. When vehicles entered the signalized intersections at the end of “green wave,” they might execute the unsafe behaviors of speeding and frequent lane changing [27]. In addition, there was a longer vehicular delay on the adjacent uncoordinated arterials, where the drivers became inpatient and might engage in “road rage” [8].

In conclusion, the safety study of signal coordination deserves an in-depth exploration with effective evaluation models.

### 2.3. Statistical Models of Crash Frequency considering the Heterogeneous Correlations.

There were a number of classical safety evaluation models of crash frequency, such as the Poisson regression model (the primary Bayesian counting model) [28], negative binomial/Poission-Gamma regression model (solving the issue of overdispersion) [29], and zero-inflated Poisson regression model (considering the existence of zero crashes in units) [6]. Based on the traditional approaches, several models revealed the heterogeneity to ensure the stability of regression results, including the random effects model, finite mixture/latent class model, Markov switching model, and generalized linear model [30]. Alternatively, the random effects model is more flexible to reflect the heterogeneous differentiation as the random term [30, 31]. With the method of maximum likelihood estimation, the model could acquire a better goodness-of-fit for variables.

#### 2.3.1. Multivariate Heterogeneity.

Crashes of different types may show the diverse contributors [32, 33]. For achieving the safety analysis at the microscopic aspect, crashes are usually classified into several levels as injury, PDO, and fatality. The multivariate random effects model is able to reflect heterogeneous correlations among crashes with different severities [30, 34].

#### 2.3.2. Structural Heterogeneity.

The structural heterogeneity exists between crashes at units with different road design and traffic flow features (e.g., roadway segments and signalized intersections along the coordinated arterials) [9]. Scholars concluded that the random effects model suffices to address the structural heterogeneity [30, 35]. However, the structural relationship between the units was not truly formulated. With the joint random effects model, Fan et al. [9] recently rebuilt the structural association between the roadway segments and signalized intersections.

#### 2.3.3. Spatiotemporal Heterogeneity.

In contrast to the structural correlation, spatial heterogeneity derives from crashes in units with similar characteristics [36, 37]. The analysis of spatial heterogeneity has attracted a wide popularity in the aspects of intersections [38], segments [36, 39], TAZs [33, 40], corridors [41], counties [26], and states/provinces [42]. Generally, the spatial heterogeneity in the random effects model was reflected as the distance function of homogeneous units [10].

There are the analyses of temporal variation among multiple years of crash data. For instance, Cheng et al. [43] explored the spatial and temporal heterogeneities (i.e., spatiotemporal heterogeneity) in the same model, where the specific forms included the fixed spatial and linear temporal association, fixed spatial and linear-quadratic temporal association, and time-varying spatial association.

### 3. Data

#### 3.1. Description of Signal-Coordinated Arterials.

The data records six arterials with the signal coordination from Ann Arbor City (Michigan State, USA) [44], located in Eisenhower, Packard, Plymouth, Stadium, State, and Washtenaw Roads. Since 2009, the signal-coordinated arterials apply the Adaptive Control System (ATCS), which poses the advantages of presenting the flexible signal plan, reflecting the timely traffic status, and providing the reasonable interference of signal coordination [45]. More specifically, ATCS detects the real-time traffic status (i.e., occupancy, flow, and congestion) with the use of monitors installed upstream of signalized intersections. The detected data serves to predict the vehicular queuing length, thus optimizing the corresponding signal attributes. It is mentioned that the relative signal attributes (cycle, timing, and phase) could not be directly acquired from the arterials due to their changeable features with the time variation.

Along the coordinated arterials, we investigated 38 roadway segments and 38 signalized intersections. For the purpose of differentiating the specific location, each signalized intersection locates in the area consisting of the central district and safety affecting region (i.e., 250 feet away...
from the stop bar) (Figure 1) [36, 39]. The roadway segment lies in the section between two adjacent intersections. Every TAZ is formulated by the signalized intersection and its adjacent roadway segment. With the distinguished units, the paper extracts the dependent variables, road design features, and traffic characteristics, respectively (Table 1).

Google Earth 2014 is capable of pinpointing the latitude and longitude intervals for the 38 roadway segments (or signalized intersections). With the geographical information, nine-year (2009–2014) crash data were retrieved from the Michigan Department of Transportation (MDOT). The total crash count, excluding the crashes related to animals, nonmotor vehicles (e.g., bicycles and motorcycles), and pedestrians, is 5613 along the coordinated arterials composed by 1996 roadway segments and 3647 crashes of signalized intersections. There are 1576/2989 PDO crashes and 387/657 injury crashes occurring on the segments/intersections. The fatal crashes are filtered due to the small sample size.

Road design variables, revealing operational requirements of the signal-coordinated arterials, are obtained by the road survey with Google Earth 2014. Meanwhile, the paper acquires the related traffic characteristics with the software of ArcGIS and the traffic volume map of the Southeast Michigan Council of Governments (SEMCOG) [46]. The related traffic variables suffice to describe the development status of arterials under the coordinated signalization.

3.2. Data Processing. According to the conclusion of Alarifi et al. [39], multicollinearity exists among the variables. The potential disturbance involved in the data may lead to unstable results during the modeling process. For alleviating the multicollinear effect, we employ the techniques of product-moment and Spearman rank-order [47]. When the correlation coefficient is larger than 0.6 between two estimated variables, only the more important one (i.e., the variable reflecting the characteristic of arterials more in detail) will remain. The filtered variables are in italics in Table 1.

4. Methodology

4.1. Model Configuration. There are seven models in the paper, including four existing models (i.e., PLN model, JMPLN model, JMPLN-CAR model, and MPLN model) and three newly formulated models (i.e., JMPLN-CAR model with fixed spatial and linear temporal variation, JMPLN-CAR model with fixed spatial and linear-quadratic temporal variation, and JMPLN-CAR model with time-varying spatial interaction). Detailed descriptions are summarized in Table 2.

According to Table 2, the comparison of models 1-2 verifies the existence of multivariate heterogeneity between PDO and injury crashes. Similarly, the paper could compare models 2-3, 3-4, and 5-7 to explore the structural, spatial, and spatiotemporal heterogeneities, respectively.

4.1.1. Poisson Log-Normal Model (Model 1). The framework of model 1 is specified as [34]

\[ y_{ij}^{k} \sim \text{Poisson}(\theta_{ij}^{k}), \]
\[ E(y_{ij}^{k}) = \theta_{ij}^{k}, \]

where \( y_{ij}^{k} \) stands for the crash count of unit \( i \) (\( i = 1, 2, \ldots, M \); \( M = 38 \) is the total number of units) for severity level \( j \) (\( j = 1, 2, \ldots, S \); \( j = 1 \) for injury and \( j = 2 \) for PDO) that belongs to structure \( k \) (\( k = 1 \) for roadway segment and \( k = 2 \) for signalized intersection). \( \theta_{ij}^{k} \) represents the expectation of \( y_{ij}^{k} \).

The probability of crash counts satisfies the formula as follows:

\[ P(y_{ij}^{k} = q) = \frac{e^{-\lambda_{ij}^{k}} \lambda_{ij}^{kq}}{q!}, \]
\[ \lambda_{ij}^{k} = \exp\left(\beta_{1}^{k} x_{ij}^{k} + D_{ij}^{k} + \mu_{ij}^{k}\right), \]

where \( \lambda_{ij}^{k} \) is the intercept of severity level \( j \) and structure \( k \); \( \beta_{1}^{k} \) reflects the regression vector related to the explanatory vector \( x_{ij}^{k} \); \( D_{ij}^{k} \) stands for the dichotomous variable, equaling to the length of segment \( i \) if \( k = 1 \) and equaling to 0 if \( k = 2 \); \( \mu_{ij}^{k} \sim \text{Normal}(0, (1/\theta^{2})) \); and \( \theta_{ij}^{k} \sim \text{gamma}(0.01, 0.01) \) [8].

4.1.2. Multivariate Poisson Log-Normal Model (Model 2). We have

\[ \log(\theta_{ij}^{k}) = \lambda_{ij}^{k} + \beta_{1}^{k} x_{ij}^{k} + D_{ij}^{k} + \epsilon_{ij}^{k}, \]

where \( \epsilon_{ij}^{k} \sim MN_{S}^{k}(0, \Sigma_{e}) \) indicates the multivariate-heterogeneous effect between injury and PDO crashes [8]. Also, \( \Sigma_{e} \) is the corresponding matrix of variance-covariance denoted by

\[ \Sigma_{e}^{-1} \sim \text{Wishart}(R, n), \]

where \( \Sigma_{e}^{-1} \) is the inverse matrix of \( \Sigma_{e} \); \( R \) represents the scale matrix of Jonathan et al. [48]; and \( n \) is the freedom degree, assigned as 2 according to the classification of structure \( k \). Model 2 further discusses the multivariate heterogeneity on roadway segments (or at signalized intersections) in the microscopic aspect compared to model 1.

4.1.3. Joint Multivariate Poisson Log-Normal Model (Model 3). Compared to model 2, model 3 incorporates the structural-heterogeneous correlation in each TAZ from the mesoscopic level. Inspired by the study of Fan et al. [9], model 3 follows the formula as

\[ \log(\theta_{ij}^{k}) = \lambda_{ij}^{k} + \beta_{1}^{k} x_{ij}^{k} + D_{ij}^{k} + \epsilon_{ij}^{k} + \eta z_{ij} + H_{ij}, \]

where \( \eta \) is the vector of regression parameters related to the explanatory vector \( z_{ij} \) of TAZ and \( H_{ij} \) represents the term of structural heterogeneity between each segment and its adjacent intersection [9]. More specifically, \( H_{1} = \delta_{1} \mu_{ij}^{k} \) is the
structural heterogeneity from the intersections acting on the segments, while $H_2 = (\delta_2 \mu_1 + \delta_3 \mu_2)$ stands for the structural heterogeneity from the segments acting on the intersections; $\delta_m$ ($m = 1, 2, 3$) reflects the estimated coefficient. $\mu_i$ satisfies $\text{Normal}(0, (1/\tau_i))$ and $\tau_i$ is the precision parameter [33].

4.1.4. Joint Multivariate Poisson Log-Normal Conditional Autoregressive Model (Model 4). Model 4 further considers the spatial-heterogeneous correlation among the TAZs. The corresponding equation is shown as

$$
\log(\theta_{ij}) = \lambda_j^k + \beta_k^i x_{ij} + D_j^k + \epsilon_{ij} + \eta_{Z_i} + H_k + \psi_i,
$$

(7)

where $\psi_i$ stands for the CAR term of spatial relationship. The specific distribution of $\psi_i$ is assigned as

$$
\psi_i | \psi_{-i} \sim \text{Normal} \left( \frac{\sum_{j \neq i} w_{ij} \psi_i}{\sum_{j \neq i} w_{ij} \sum_{j \neq i} w_{ij}}, \frac{\sigma^2}{\sum_{j \neq i} w_{ij}} \right),
$$

(8)

where $\psi_{-i}$ is the set derived from $\psi_i$ [39]; $(1/\sigma^2)$ reflects the CAR precision parameter assumed to be the gamma hyper-
where each entry $w_{ij}$ is the spatial heterogeneity between units $i$ and $j$. The formula of $w_{ij}$ is given as

$$w_{ij} = \begin{cases} e^{-\alpha d_{ij}}, & \text{if } i \text{ and } j \text{ are on the same arterial,} \\ 0, & \text{if } i \text{ and } j \text{ are on different arterials,} \end{cases}$$

where $d_{ij}$ is shown as the geometry centroid length and $\alpha (>0)$ indicates the decline rate of distance [49].

Derived from model 4, models 5–7 identify the different types of spatiotemporal-heterogeneous correlations in Sections 4.1.5–4.1.7, respectively.

### 4.1.5. Joint Multivariate Poisson Log-Normal Conditional Autoregressive Model with Fixed Spatial and Linear Temporal Variation (Model 5)

$$\log(\hat{\theta}_{ij}) = \lambda_j + \beta_i^k X_{ik} + D_k^t + \epsilon_{ij} + \eta Z_i + H_k + \tilde{ST}_t^i,$$

$$\tilde{ST}_t^i = \psi_i + (\xi + \delta_t) \cdot t,$$

where $\hat{\theta}_{ij}$ is the crash count of $\theta_{ij}$ divided by time $t$ ($t = 1, 2, \ldots, T$); $\tilde{ST}_t^i = \psi_i + (\xi + \delta_t) \cdot t$ reveals the spatiotemporal heterogeneity, consisting of the fixed spatial correlation $\psi_i$ and the linear temporal variation $(\xi + \delta_t) \cdot t$; $t$ is the linear trend of time; $\xi$ is the average variation among all units; and $\delta_t$ follows the similar distribution as $\psi_i$.

In model 5, it is emphasized that the spatial variation ($\psi_i$ and $\delta_t$) is independent of the temporal changing ($t$) [43].

### 4.1.6. Joint Multivariate Poisson Log-Normal Conditional Autoregressive Model with Fixed Spatial and Linear-Quadratic Temporal Variation (Model 6). We have

$$\log(\hat{\theta}_{ij}) = \lambda_j + \beta_i^k X_{ik} + D_k^t + \epsilon_{ij} + \eta Z_i + H_k + \tilde{ST}_t^i,$$

$$\tilde{ST}_t^i = \psi_i + (\xi + \delta_t) \cdot t + \gamma \cdot t^2.$$

Model 6 addresses the fixed spatial and linear-quadratic temporal variation ($\tilde{ST}_t^i$) based on model 5. In equation (12), $\gamma \cdot t^2$ reflects the quadratic temporal variation; and $\gamma$ is the coefficient corresponding to the quadratic time $t^2$.

$\psi_i$ and $\delta_t$ are unrelated to $t$ and $t^2$, which indicates that the temporal trend does not affect the spatial connectivity and vice versa [10].

### 4.1.7. Joint Multivariate Poisson Log-Normal Conditional Autoregressive Model with Time-Varying Spatial Interaction (Model 7). Model 7 presents the spatiotemporal interaction between the spatial and temporal heterogeneity in reality [10]:

$$\log(\hat{\theta}_{ij}) = \lambda_j + \beta_i^k X_{ik} + D_k^t + \epsilon_{ij} + \eta Z_i + H_k + \tilde{ST}_t^i,$$

$$\tilde{ST}_t^i = \psi_i,$$

where the time-varying spatial interaction $\psi_i$ satisfies the distribution as $MN(0, \Sigma_k)$. $\Sigma_k$ is the variance-covariance matrix expressed as

$$\Sigma_k^{-1} \sim \text{ Wishart}(\bar{R}, \mathbf{n}),$$

$\Sigma_k^{-1}$ is the inverse matrix of $\Sigma_k$; $\bar{R}$ is the scale matrix defined by Aguero-Valverde and Jovanis [50]; and $\mathbf{n}$ stands for the degree of freedom [10].

### 4.2. Bayesian Inference and Model Comparison. Three types of methods are feasible to identify the differentiation among the models, including the classical criteria (i.e., WMAD/MAD and WMSPE/MSPE), pervasive criteria (i.e., TDIC/DIC), and emerging criteria (i.e., TLPML/LPML). The summary of the criteria is presented in Table 3.

#### 4.2.1. Mean Absolute Deviance and Mean Squared Prediction Error. The classical approaches of MAD and MSPE are employed to compare the model fitness [41, 52, 53]. Regarding the structure $k$, they are separately presented in the following equations:

<table>
<thead>
<tr>
<th>Model</th>
<th>Abbreviation</th>
<th>Section</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLN model</td>
<td>Model 1</td>
<td>4.1.1</td>
<td>Foundation of the crash frequency analysis with the random effect</td>
</tr>
<tr>
<td>MPLN model</td>
<td>Model 2</td>
<td>4.1.2</td>
<td>Identifies the multivariate heterogeneity</td>
</tr>
<tr>
<td>JMPLN model</td>
<td>Model 3</td>
<td>4.1.3</td>
<td>Explores the multivariate and structural heterogeneities</td>
</tr>
<tr>
<td>JMPLN-CAR model</td>
<td>Model 4</td>
<td>4.1.4</td>
<td>Captures the multivariate, structural, and spatial heterogeneities</td>
</tr>
<tr>
<td>JMPLN-CAR model with fixed spatial and linear temporal variation</td>
<td>Model 5</td>
<td>4.1.5</td>
<td>Discusses the multivariate, structural, and spatial heterogeneities</td>
</tr>
<tr>
<td>JMPLN-CAR model with fixed spatial and linear-quadratic temporal variation</td>
<td>Model 6</td>
<td>4.1.6</td>
<td>Addresses the multivariate, structural, and spatiotemporal heterogeneities</td>
</tr>
<tr>
<td>JMPLN-CAR model with time-varying spatial interaction</td>
<td>Model 7</td>
<td>4.1.7</td>
<td>Reflects the multivariate, structural, and spatiotemporal heterogeneities</td>
</tr>
</tbody>
</table>

Note: The full name of the models is shown in Section 1.
shown as [52].

Note. The full name of the criteria is shown in Section 1; the detailed description of the above criteria is in Fan et al. [10], Heydari et al. [51], and Zeng and Huang [52].

4.2.2. Deviance Information Criterion. The formula of DIC for structure $k$ is described as [36, 39, 40, 53, 54]

$$ \text{DIC}^k = \text{TDIC}^k + \text{MSPE}^k, $$

where $\text{TDIC}^k$ represents the posterior mean deviance and $\text{MSPE}^k$ shows the complexity measurement of the model.

Total deviance information criterion (TDIC) of total TAZs is formulated as

$$ \text{TDIC} = \text{TDIC}^1 + \text{TDIC}^2. $$

4.2.3. Log Pseudo Marginal Likelihoods. LPML of each structure $k$ satisfies the formula as follows [10, 43, 51]:

$$ \text{LPML}^k = \log \prod_{i=1}^{M} \prod_{j=1}^{S} \text{CPO}_{ij}^k, $$

where $\text{CPO}_{ij}^k$ is the conditional predictive ordinate given by

$$ \text{CPO}_{ij}^k = \left( \frac{1}{T} \sum_{t=1}^{T} \frac{1}{f(t)} \right)^{-1}. $$

In equation (21), $f(\cdot | \cdot)$ is the conditional predictive distribution.

In terms of TAZs, the total LPML is denoted as follows:

$$ \text{TLPLM} = \text{LPML}^1 + \text{LPML}^2 = \log \prod_{i=1}^{M} \prod_{j=1}^{S} (\text{CPO}_{ij}^1 \text{CPO}_{ij}^2). $$

4.3. Variation Proportion. The paper utilizes $\alpha_1$, $\alpha_2$, and $\alpha_3$ to reveal the variation proportion explained by the multivariate, structural, and spatiotemporal correlations (i.e., $\varepsilon$, $H$, and $ST$), respectively. The specific formula is shown as

$$ \alpha_1 = \frac{\text{sd}(\varepsilon)}{\text{sd}(\varepsilon) + \text{sd}(H) + \text{sd}(ST)}, $$

$$ \alpha_2 = \frac{\text{sd}(H)}{\text{sd}(\varepsilon) + \text{sd}(H) + \text{sd}(ST)}, $$

$$ \alpha_3 = \frac{\text{sd}(ST)}{\text{sd}(\varepsilon) + \text{sd}(H) + \text{sd}(ST)}, $$

where $\text{sd}(\cdot)$ is the standard deviation. The lower value of $\alpha_m$ ($m = 1, 2, 3$) indicates the relatively weaker heterogeneous correlation corresponding to the total sum of heterogeneities [34].

5. Results

The paper conducts the modeling estimation with the software of WinBUGs 1.4.3 [55], where the total iterations perform as 30,000 containing discarded 10,000 iterations at the burn-in-step.

5.1. Goodness-of-Fit for Models. Tables 4 and 5 provide the predictive performance of models 1–4 and 5–7, respectively. According to Table 4, model 4 with the lowest TDIC value (888.565) prefers models 1–3 (separately with the TDIC value of 903.119, 926.193, and 973.746). However, the evaluation result of WMAD and WMSPE is inconsistent with the counterpart of TDIC. It is because WMAD and
WMSPE fail to account for the complexity of models in contrast to TDIC (Table 3). Table 5 adds the approach of TLPML to further quantify the division of temporal dimension in models 5–7. Model 5 notably rules out models 6–7 as associating with the highest TLPML value (−65.750 > −66.890 > −75.470) and the lowest TDIC value (4780.790 < 4785.050 < 4940.530). However, the estimate of WMAD and WMSPE generates the biases compared to TLPML and TDIC according to Table 3. With the above criteria, it is difficult to directly compare the fitness between models 1–4 and 5–7 due to the increase of temporal dimension in latter models. Therefore, the paper utilizes the method of variation proportion to complete the model comparison (Table 6). It is shown that the values of sd(ε), sd(H), and sd(ST) are greater than 0, which verifies the existence of multivariate, structural, and spatiotemporal heterogeneities. By incorporating the multiple-heterogeneous correlations simultaneously, models 5–7 implicitly obtain a better predictive performance compared to models 1–4. Combined with the modeling results in Table 5, it can be summarized that model 5 poses a better fitness than other models. The parameter estimation and discussions below are based on model 5.

Meanwhile, the value of α₁ (0.764) is larger than α₄ (0.166) and α₂ (0.070) (Table 6). It implies that multivariate heterogeneity has the most powerful impact on the assessment results, followed by the spatiotemporal and structural heterogeneities [34, 40, 53].

### 5.2. Parameter Estimation

Presented in Table 7 is the regression result of models 5–7. The corresponding results of models 1–2 and 3–4 are shown in the Supplementary Materials section in Tables S1 and S2, respectively. Meanwhile, the point range plots of Table 7 are provided in the Supplementary Materials section in Figures S1–S5. Based on the results of roadway segments in model 5, several contributors are positively associated with the occurrence of injury and PDO crashes, including the median opening width, lane density, posted speed limit, and log AADT. Meanwhile, the principal arterial is assumed to cause more crashes than the minor arterial according to the modeling result of functional roadway classification. In terms of the TAZs, the estimation results are summarized as (1) the increase of average AADT causing more crashes; (2) the higher average posted speed limit decreasing the possibility of crash occurrence; and (3) the length of TAZ not being a significant indicator of crash frequency. Regarding the signalized intersections, four variables (i.e., NETMA, NETMI, number of legs, and log AADT on major road) are the positively statistical evidence impacting on the number of injury and PDO crashes at the significance level of 95%. PLMA, PLMI, PMMA, and PMMI

### Table 4: Summarized results of the predictive performance for models 1–4.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIC</td>
<td>433.301</td>
<td>973.746</td>
<td>399.828</td>
<td>526.365</td>
</tr>
<tr>
<td>TDIC</td>
<td>973.746</td>
<td>399.828</td>
<td>526.365</td>
<td>399.828</td>
</tr>
<tr>
<td>MAD</td>
<td>3.236</td>
<td>4.886</td>
<td>3.030</td>
<td>4.769</td>
</tr>
<tr>
<td>WMAD</td>
<td>26.030</td>
<td>48.310</td>
<td>25.060</td>
<td>46.840</td>
</tr>
<tr>
<td>WMSPE</td>
<td>37.170</td>
<td>35.950</td>
<td>35.800</td>
<td>36.320</td>
</tr>
</tbody>
</table>

**Note.** From model 3, the roadway segment and signalized intersection are incorporated in the whole TAZ (Section 4.1.3).

### Table 5: Summarized results of the predictive performance for models 5–7.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Model 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLPML</td>
<td>−65.750</td>
<td>−75.470</td>
<td>−69.800</td>
</tr>
<tr>
<td>TDIC</td>
<td>4780.790</td>
<td>4785.050</td>
<td>4785.050</td>
</tr>
<tr>
<td>WMAD</td>
<td>1.473</td>
<td>1.465</td>
<td>1.592</td>
</tr>
<tr>
<td>WMSPE</td>
<td>5.379</td>
<td>5.292</td>
<td>6.871</td>
</tr>
</tbody>
</table>

### Table 6: Summary of the variation proportion related to models 5–7.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Model 5 (95% BCI)</th>
<th>Model 6 (95% BCI)</th>
<th>Model 7 (95% BCI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sd(ε)</td>
<td>5.179 (4.161, 6.748)</td>
<td>43.360 (20.060, 66.430)</td>
<td>4.996 (3.843, 6.169)</td>
</tr>
<tr>
<td>sd(H)</td>
<td>0.481 (0.098, 1.172)</td>
<td>0.568 (0.096, 1.705)</td>
<td>0.347 (0.065, 1.003)</td>
</tr>
<tr>
<td>sd(ST)</td>
<td>1.126 (0.556, 1.807)</td>
<td>11.580 (0.905, 19.810)</td>
<td>1.215 (0.987, 1.392)</td>
</tr>
<tr>
<td>α₁</td>
<td>0.076 (0.015, 0.164)</td>
<td>0.079 (0.002, 0.028)</td>
<td>0.052 (0.011, 0.138)</td>
</tr>
<tr>
<td>α₂</td>
<td>0.166 (0.089, 0.258)</td>
<td>0.193 (0.043, 0.239)</td>
<td>0.187 (0.151, 0.236)</td>
</tr>
<tr>
<td>α₃</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** BCI is the Bayesian "credible" interval.
### Table 7: Regressive results of variables related to models 5–7.

<table>
<thead>
<tr>
<th>Location</th>
<th>Type</th>
<th>Variable</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Model 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (SD)</td>
<td>95% BCI</td>
<td>Mean (SD)</td>
<td>95% BCI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intercept</td>
<td>0.998 (0.033)</td>
<td>0.999 (0.033)</td>
<td>0.998 (0.031)</td>
</tr>
<tr>
<td>Injury crash</td>
<td></td>
<td>Median opening width</td>
<td>0.492 (0.157)</td>
<td>0.862 (0.207)</td>
<td>0.815 (0.221)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Functional roadway classification</td>
<td>0.972 (0.195)</td>
<td>0.995 (0.194)</td>
<td>0.986 (0.190)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lane density</td>
<td>0.655 (0.130)</td>
<td>0.839 (0.150)</td>
<td>0.615 (0.167)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posted speed limit</td>
<td>0.656 (0.183)</td>
<td>0.842 (0.231)</td>
<td>0.664 (0.210)</td>
</tr>
<tr>
<td>Roadway segment</td>
<td></td>
<td>Log AADT</td>
<td>0.701 (0.161)</td>
<td>0.820 (0.205)</td>
<td>0.750 (0.168)</td>
</tr>
<tr>
<td>PDO crash</td>
<td></td>
<td>Intercept</td>
<td>0.997 (0.035)</td>
<td>0.999 (0.036)</td>
<td>0.997 (0.032)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Median width</td>
<td>0.496 (0.156)</td>
<td>0.878 (0.205)</td>
<td>0.432 (0.229)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Functional roadway classification</td>
<td>0.953 (0.193)</td>
<td>0.994 (0.200)</td>
<td>0.986 (0.188)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lane density</td>
<td>0.749 (0.126)</td>
<td>0.940 (0.152)</td>
<td>0.723 (0.174)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posted speed limit</td>
<td>0.635 (0.262)</td>
<td>0.573 (0.266)</td>
<td>0.695 (0.208)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Log AADT</td>
<td>0.652 (0.174)</td>
<td>0.728 (0.162)</td>
<td>0.700 (0.151)</td>
</tr>
<tr>
<td>TAZ Crash</td>
<td></td>
<td>Length</td>
<td>-0.170 (0.119)</td>
<td>0.430 (0.224)</td>
<td>-0.381 (0.142)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Average posted speed limit</td>
<td>-0.586 (0.074)</td>
<td>0.576 (0.177)</td>
<td>-0.395 (0.085)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Average AADT</td>
<td>0.784 (0.138)</td>
<td>0.960 (0.162)</td>
<td>0.825 (0.165)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intercept</td>
<td>0.822 (0.231)</td>
<td>0.845 (0.234)</td>
<td>0.878 (0.214)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NETMA</td>
<td>0.551 (0.205)</td>
<td>0.583 (0.193)</td>
<td>0.580 (0.175)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NETMI</td>
<td>0.804 (0.161)</td>
<td>0.741 (0.171)</td>
<td>0.586 (0.156)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLMA</td>
<td>0.884 (0.178)</td>
<td>0.946 (0.221)</td>
<td>0.911 (0.170)</td>
</tr>
<tr>
<td>Injury crash</td>
<td></td>
<td>PLMI</td>
<td>0.975 (0.176)</td>
<td>0.452 (0.133)</td>
<td>0.643 (0.137)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of legs</td>
<td>0.832 (0.19)</td>
<td>0.876 (0.226)</td>
<td>0.888 (0.180)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PMMA</td>
<td>0.887 (0.186)</td>
<td>0.950 (0.232)</td>
<td>0.912 (0.191)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PMMI</td>
<td>0.846 (0.179)</td>
<td>0.943 (0.226)</td>
<td>0.897 (0.193)</td>
</tr>
<tr>
<td></td>
<td>ASD</td>
<td>-0.097 (0.214)</td>
<td>-0.553 (0.179)</td>
<td>-0.194 (0.121)</td>
<td></td>
</tr>
<tr>
<td>Signalized intersection</td>
<td></td>
<td>Log AADT on major road</td>
<td>0.536 (0.133)</td>
<td>0.553 (0.223)</td>
<td>0.699 (0.152)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intercept</td>
<td>0.938 (0.250)</td>
<td>0.924 (0.264)</td>
<td>0.970 (0.199)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NETMA</td>
<td>0.904 (0.21)</td>
<td>0.767 (0.239)</td>
<td>0.648 (0.123)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NETMI</td>
<td>0.900 (0.177)</td>
<td>0.838 (0.207)</td>
<td>0.679 (0.169)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLMA</td>
<td>1.018 (0.195)</td>
<td>0.988 (0.243)</td>
<td>0.878 (0.169)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PMMA</td>
<td>0.892 (0.189)</td>
<td>0.952 (0.231)</td>
<td>0.826 (0.165)</td>
</tr>
<tr>
<td>PDO crash</td>
<td></td>
<td>Number of legs</td>
<td>0.961 (0.200)</td>
<td>0.903 (0.237)</td>
<td>0.941 (0.176)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PMMA</td>
<td>1.041 (0.203)</td>
<td>1.001 (0.229)</td>
<td>0.839 (0.179)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PMMI</td>
<td>1.039 (0.195)</td>
<td>0.964 (0.237)</td>
<td>0.916 (0.179)</td>
</tr>
<tr>
<td></td>
<td>ASD</td>
<td>-0.115 (0.271)</td>
<td>-0.756 (0.239)</td>
<td>-1.275 (0.654)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Log AADT on major road</td>
<td>0.579 (0.140)</td>
<td>0.219 (0.377)</td>
<td>0.048 (0.089)</td>
</tr>
</tbody>
</table>

Note. *Variable is positive significant at 92.5% BCI; †variable is positive significant at 47.5% BCI; ‡variable is negative significant at 47.5% BCI; §variable is negative significant at 72.5% BCI; ††variable is positive significant at 92.5% BCI.
increase the crash count. The variation of ASD is unrelated to both crash occurrences.

6. Conclusion and Discussion

Previous research lacks the comprehensive heterogeneity consideration for the signal-coordinated arterials. This paper aims to build the safety evaluation model by incorporating multiple heterogeneities simultaneously, including the multivariate heterogeneity (at the microscopic level), structural heterogeneity (from the mesoscopic aspect), and spatiotemporal heterogeneity (in the macroscopic viewpoint). The model serves to identify the significant contributors to crash frequency, thus achieving the accurate safety analysis of arterials with the coordinated signalization.

Tables 5 and 6 show that models 5–7 perform better than models 1–4 due to the existence of multivariate, structural, and spatiotemporal heterogeneities. Signal coordination forms the "green wave" to impose the high-speed limit for the consecutively vehicular platoon, where the vehicles are densely correlated. Thus, the coupled vehicles connect as a whole generating the multivariate, structural, and spatiotemporal heterogeneities. The conclusion verifies the necessity of proposing the safety evaluation model with the multiple heterogeneities. In addition, it is imperative to provide the safety improvement countermeasures to alleviate the correlated crashes caused by the multiple heterogeneities, such as (1) adding the content of signal-coordinated arterials during the driving license test to enhance the drivers' theoretical and practical cognition; (2) setting the warning signs at the entrance of coordinated arterials [9]; (3) equipping the variable message sign system (VMSS) to present the traffic information for vehicles in real time [56]; (4) providing vehicles with the advanced technologies of autonomous driving, connected vehicle, and artificial intelligence [57–59], thus achieving the accurate crash control and eliminating the interference of humans.

Model 5 is superior to models 6-7 (Table 5), implying that there exists the fixed spatial and linear temporal variation (Section 4.1.5). Specifically, it indicates that the variation of spatial heterogeneity is unrelated to the temporal changing among the TAZs. It is possible that, with the aid of signal coordination, the stable spatial correlations on signalized intersections evenly affect the whole TAZs [8]. The signal coordination allows the vehicular platoon to pass through the intersections with the invariant headways, while the invariant headways are associated with the stable distribution of platoon, resulting in the unchanged spatial heterogeneity. Given the consistent spatial association, it is necessary to divide the subcontrol traffic zones based on the correlative degree of vehicles in real time. With the response to the dynamic traffic demand, the approach can avoid the "green wave" being too long or too short, thus ensuring the stability of vehicular movement and alleviating the unsafe behaviors (e.g., acceleration and "stop and go") [60, 61].

Meanwhile, it is found that multivariate heterogeneity obtains a stronger impact on the evaluation results than the structural and spatiotemporal heterogeneities (Table 6). In view of the differentiation among the multivariate, structural, and spatiotemporal heterogeneities, appropriate suggestions are recommended for divers, such as (1) paying more attention to the vehicular distribution in real time, thus dynamically responding to the correlative criterion; (2) maintaining the driving speed and headway to ensure the awareness of safety risks; (3) adopting the actuated advance warning system (ADWS) [62]; and (4) equipping the onboard sensors in vehicles to retrieve a large amount of data (e.g., the acceleration, application status of vehicular components, and the distance between other vehicles) with the video recording technology [63].

For the roadway segments (Table 7), median opening width is positively correlated with injury and PDO crash counts. A feasible explanation is that the high-speed vehicular platoon migrates from the signalized intersections to the adjacent segments due to the implementation of signal coordination. Without the signalized control, the aggregated vehicles on the segments pose relatively short headways [8, 9]. Under the scenario, the wider median opening may incur more risky driving behaviors (e.g., the simultaneous U-turn of vehicles), thus resulting in the phenomenon of traffic jams more easily. Similarly, the increase of lane density, posted speed limit, log AADT, and functional roadway classification causes safety hazards along the roadway segments [39, 64–66]. In order to alleviate the crash hazards of roadway segments along the coordinated arterials, the countermeasures are provided: (1) reducing the width of median opening, evenly decreasing the setting of median opening; (2) implementing the low-speed limit on the segments; (3) applying the vibrating deceleration pavement markings to control the vehicular speed; and (4) maintaining a sufficient sight distance for the drivers.

At the signalized intersections (Table 7), the higher possibility of injury and PDO crash occurrences is due to the increase of NETMA, NETMI, the number of legs, and log AADT on the major road. Meanwhile, PLMA, PLMI, PMMA, and PMMI result in more crash counts. The above findings are in accordance with the previous literature [8, 39, 67–69]. The variable of ASD is independent of crash count. As opposed to the minor road, signal coordination allocates the adequate "green wave" on the major road with a high-speed limit [9, 41]. In the presentation of decreasing the ASD (or appropriately increasing the posted speed limit on the minor road), the crash variation of whole intersections is insignificant due to the limited traffic flow and "green bandwidth" on the minor road.

Although the study considers the potential disturbance involved in the data, the issue of multicollinearity may not be completely resolved. The safety evaluation of the signal-coordinated system needs to be further explored by employing more effective methods to address the multicollinearity among variables, e.g., testing each variable in the safety evaluation model and then excluding ones to be highly correlated.

Data Availability

The original accident data used to support the findings of this study are currently under embargo. Requests for data, 12
months after publication of this article, will be considered by the corresponding author.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Acknowledgments
The work was supported by the National Natural Science Foundation of China (NSFC-71771191), the Doctoral Research Foundation of Taiyuan University of Science and Technology (20202048), the Shanxi Scholarship Council of China (2021-136), and the Natural Science Foundation of Zhejiang Province (LQ21G010002). The authors would like to thank He Huang (School of Transportation and Logistics Engineering, Taiyuan University of Science and Technology) for the generous help.

Supplementary Materials
The following is the concise description for supplementary files. Tables S1 and S2 present the regression results of -Q_he following is the concise description for supplementary files. Meanwhile, Figures S1–S5 provide the point range plots of Table 7 from Section 5.2. (Supplementary Materials)

References


