

Article

# Exploring Spatiotemporal Patterns of Expressway Traffic Accidents Based on Density Clustering and Bayesian Network

Yunfei Zhang <sup>1,2,\*</sup> , Fangqi Zhu <sup>2</sup>, Qiuping Li <sup>3</sup> , Zehang Qiu <sup>2</sup> and Yajun Xie <sup>2</sup>

<sup>1</sup> Engineering Laboratory of Spatial Information Technology of Highway Geological Disaster Early Warning in Hunan Province, Changsha University of Science & Technology, Changsha 410114, China

<sup>2</sup> School of Traffic & Transportation Engineering, Changsha University of Science & Technology, Changsha 410114, China

<sup>3</sup> School of Geography and Planning, Sun Yat-sen University, Guangzhou 510006, China

\* Correspondence: zhang.yunfei@csust.edu.cn

**Abstract:** Exploring spatiotemporal patterns of traffic accidents from historic crash databases is one essential prerequisite for road safety management and traffic risk prevention. Presently, with the emergence of GIS and data mining technologies, numerous geospatial analysis methods have been successfully adopted for traffic accident analysis. As characterized by high driving speeds, diverse vehicle types, and isolated traffic environments, expressways are confronted with more serious accident risks than urban roads. In this paper, we propose a combined method based on improved density clustering and the Bayesian inference network to explore spatiotemporal patterns of expressway accidents. Firstly, the spatiotemporal accident neighborhood is integrated into the DBSCAN clustering algorithm to discover multi-scale expressway black spots. Secondly, the Bayesian network model is separately employed in both local-scale black spots and regional-scale expressway networks to fully explore spatially heterogeneous accident factors in various black spots and expressways. The experimental results show that the proposed method can correctly extract spatiotemporal aggregation patterns of multi-scale expressway black spots and meanwhile efficiently discover diverse causal factors for various black spots and expressways, providing a comprehensive analysis of accident prevention and safety management.

**Keywords:** traffic accident; black spot; causal analysis; density clustering; Bayesian network



**Citation:** Zhang, Y.; Zhu, F.; Li, Q.; Qiu, Z.; Xie, Y. Exploring Spatiotemporal Patterns of Expressway Traffic Accidents Based on Density Clustering and Bayesian Network. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 73. <https://doi.org/10.3390/ijgi12020073>

Academic Editors: Wolfgang Kainz and Hartwig H. Hochmair

Received: 16 December 2022

Revised: 11 February 2023

Accepted: 16 February 2023

Published: 18 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the last decades, road safety has been one global challenge issue that greatly threatens human life, economic property, and social consequences [1]. According to the World Health Organization (WHO), traffic accidents have always been the leading cause of death for children and young adults aged 5–29 years [2]. Moreover, it is also reported that low- and middle-income countries account for approximately 93% of accident fatalities in the world, although they own only 60% of the world's vehicles. Particularly, according to the China Statistical Yearbook [3], 244.6 thousand accidents took place in China in 2020, which lead to about 61.7 thousand deaths and 250.7 thousand injuries. Not only do accidents cause injuries and damage but they may also result in traffic congestion and even trigger local traffic paralysis [4]. Hence, it has always been a hot topic for traffic researchers and management departments to identify the accident-prone black spots and mine their spatiotemporal patterns, so as to make targeted solutions for traffic risk precaution [5,6].

Compared to urban roads, expressway traffic has the characteristics of high driving speeds, diverse vehicle types, and isolated traffic environments, and expressways may cause serious traffic accidents and trigger off massive traffic jams [7]. Presently, with the gradual improvement of China's highway infrastructures, the number and total mileage of China's expressways increase year by year, forming an efficient and dense expressway network. However, the accident rate, mortality rate and injury rate per 100 km happening

on expressways reaches 3.0, 5.1 and 3.8 times than that of ordinary roads in 2015 according to a statistical report from Traffic Management Bureau, Ministry of Public Security of China [8]. A traffic accident is a typical kind of spatiotemporal event, and historic crash databases can provide important clues for understanding the spatiotemporal scenarios of where and when traffic accidents happened. Exploring the spatiotemporal patterns of traffic accident data is one essential prerequisite for discovering expressway safety deficiencies and then making valid countermeasures. Hence, the paper concentrated on presenting a combined method based on improved density clustering and Bayesian inference network to explore the spatiotemporal patterns of traffic accidents to provide a fundamental analysis of decision making for traffic safety management.

The remainder of this paper is organized as follows. Section 2 conducts a systemic literatures review about black spot identification and accident causal analysis. Section 3 subsequently presents a methodological framework to profile spatiotemporal patterns of expressway accidents based on density clustering and the Bayesian network. In Section 4, a comprehensive analysis of the expressway accident datasets in the Hunan province, China, is undertaken to demonstrate the performance of the proposed method and interpret the spatiotemporal characteristics of Hunan expressway accidents. Finally, the conclusions and future works are discussed at the end of the paper.

## 2. Literature Review

In the past decades, numerous studies have been devoted to traffic accident analysis, which can be divided into two aspects. One aspect is identifying the accident-prone black spots (it is called an accident hotspot in urban areas) and the other one is analyzing the potential causal factors that lead to traffic accidents.

Generally, accident-prone black spots (or accident hotspots) can be defined as high-risk locations, sections, or zones where traffic accidents are more likely to happen than other places during a long-observed period in terms of accident volumes or accident characteristics [9]. Nevertheless, for various countries and regions, there is no widely accepted standard to demarcate the spatial range, time coverage, and factor significance for determining accident-prone black spots [10]. In the early stage, researchers mainly investigate various risk indicators to identify accident-prone black spots, of which the representative ones are accident count (AC) [11], accident rate (AR) [12], accident frequency (AF) [13,14], and accident density (AD) [15,16]. Particularly, some researchers attempt to split roads into geometry-homogenous segments and then introduce a hazard ratio to assess the traffic risk of these homogenous segments [17]. In view of the heterogeneities between different observation periods, some studies develop statistical models for crash risk assessment, typically such as empirical Bayesian frameworks [18] and Poisson-Tweedie models [19]. Presently, with the emergence of GIS and data mining technologies, some studies start to introduce geospatial statistical indicators [20,21] and spatial clustering methods [22,23] to find accident-prone black spots from historic crash data. For example, Blazquez et al. [24] computed Moran's I and Getis-Ord  $G_i^*$  to identify accident-prone black spots. Harirforth and Bellalite [25] combine network kernel density estimation (KDE) and network screening methods based on critical collision rates for finding accident-prone black spots. Holmgren et al. [26] applied iterative K-means clustering and DBSCAN clustering to discover the unsafe positions for urban cyclists. Those studies made significant contributions to crash data analysis, but there are still some problems to be urgently solved. On the one hand, the Modifiable Areal Unit Problem (MAUP), i.e., the optimal segmentation unit, is still an intractable problem for identifying accident-prone black spots [27]. On the other hand, the time duration and temporal periodicity of crashes are not fully considered to explore the spatiotemporal patterns of accident-prone black spots [28].

In addition to accident-prone black spot identification, the other studies mainly focus on causal analyses of traffic accidents. On the one hand, the occurrence of a road accident is possibly caused by some spatial and temporal factors at the accident location [29]. On the other hand, a sudden accident may also have an impact on a road's traffic flow [30].

Generally, the road risk factors could be summarized as four major aspects, including driver factors (e.g., speeding, drunk driving, drowsy driving), vehicle factors (e.g., brake failure), road factors (e.g., sharp turn, steep slope) and environmental factors (e.g., snow, fog, complex terrain) [31]. In the early stage, many researchers focus on one or several causal factors analysis by statistical methods. For example, Thiffault and Bergeron [32] utilized the driving simulator and proposed a steering wheel movement analysis to evaluate the impact of road environmental monotony on driver fatigue and accident risk. For another, Pervez et al. [33] adopted a seven-zone analytic approach to discuss the crash causal features for different zones of expressway long tunnels. In recent years, multi-factor analysis methods are introduced into causal analyses of traffic accidents, such as Geographically Weighted Regression (GWR) [34], association rule mining [35], and the Bayesian network [36,37]. Traffic accidents are widely accepted as spatiotemporal events, so it is indispensable to take both time and space dimensions into consideration when analyzing expressway traffic accidents. Previous studies conducted good attempts to analyze the regional-scale causal factors of traffic accidents, but regional-scale causal analysis may ignore local spatiotemporal variations of causal factors that may exist in different black spots and expressways. For example, some black spots in the expressway network may be mainly caused by complex road environments and some other ones may be greatly impacted by heavy traffic flow or unreasonable road design. Hence, it is insufficient to conduct a regional-scale causal analysis for the whole expressway but needs a divide-and-conquer strategy to study the correlations and heterogeneities of accident causal factors for various black spots and expressways.

### 3. Materials and Methods

In this paper, we propose a Spatiotemporal Density-based Accident Clustering (STDAC) method for identifying expressway accident-prone black spots and then perform a divide-and-conquer causal analysis based on Bayesian network to profile the spatiotemporal patterns of local-scale black spots and regional-scale expressway network. The main contributions of the paper can be summarized as follows. Firstly, the temporal periodicity and the accident severity are introduced to improve DBSCAN clustering algorithm, in order to discover expressway black spots with different scales. Secondly, we employ a divide-and-conquer strategy based on sophisticated Bayesian network (BN) model to explore the accident causal factors for both local-scale black spots and regional-scale expressway networks, so as to provide decision-making support for traffic safety management.

#### 3.1. Study Area and Data Preprocessing

In this study, we use the expressway accident data provided by the traffic management department of Hunan province in China. Hunan is located in the central south of China, mainly in complex terrain of mountains and hills. The study area enjoys a subtropical humid monsoon climate and are characteristic of four distinct seasons, sufficient heat, and concentrated precipitation. The whole province has more than 72.95 million population, more than 15.4 million motor vehicles, and nearly 7 thousand km expressway network according to 2021 Hunan Statistical Yearbook [38].

The experimental accident datasets (excel format) contain 98,295 simple accidents and 3240 general accidents reported between 2012 and 2016. As listed in Table 1, general accident data stores 26 accident-related attributes, but simple accident data only stores 18 accident-related attributes. Based on this reason, we use general accident data in accident causal analysis in Section 3.3 because of incomplete attributes recorded in simple accidents. Because each participant of one accident is stored as one record in raw accident data, we first grouped raw accident data by accident number and calculated the numbers of minor injuries, serious injuries, and deaths for each accident as new fields, and meanwhile, we set the field values of Car Type, Gender, and Driving Years as the personal information of the participant with main accident liability, respectively. Moreover, the accident location is recorded through a text description of the accident place and a numerical value of mileage

information. Hence, we need to calculate the spatial coordinates for each accident by matching them with the referenced expressway network. The result of geo-referenced accident data is shown in Figure 1.

**Table 1.** They include attributes in raw datasets of general accidents and simple accidents.

	Accident Attribute Name		Accident Attribute Name	
General Accident	1	Accident number	14	Weather
	2	Accident time	15	Visibility
	3	Route number	16	Illumination
	4	Road name	17	Road type
	5	Kilometers	18	Road alignment
	6	Meters	19	Gender
	7	Accident place	20	Age
	8	Accident reason	21	Traffic mode
	9	Accident situation	22	Driving years
	10	Multi-vehicles accident	23	Blood alcohol content
	11	Single vehicle accident	24	Main illegal act
	12	Carrying Hazardous goods	25	Accident liability
	13	Consequence of carrying hazardous goods	26	Injury degree
Simple Accident	1	Accident number	10	Multi-vehicles accident
	2	Accident time	11	Weather
	3	Route number	12	Road type
	4	Road name	13	Traffic mode
	5	Kilometers	14	Driving years
	6	Meters	15	Traffic mode class
	7	Accident place	16	Motor vehicle class
	8	Accident reason	17	Accident liability
	9	Accident situation	18	Car number

### 3.2. Spatiotemporal Density-Based Accident Clustering for Black Spot Identification

In the past, numerous advanced methods (e.g., KDE, spatial correlation, spatial cluster) have been proposed to identify accident-prone black spots. On the one hand, spatial statistics methods (e.g., local Moran's I and KDE) need to segment roads into homogenous sections, which brings about the Modifiable Areal Unit Problem (MAUP), that is the size of basic road unit may directly affect the precision and accuracy of black spots identification [27]. On the other hand, the existing methods mainly consider spatial aggregations of traffic accidents in the observed period, which may ignore the time evolution characteristics of accident-prone black spots [28,39].

Significantly, spatial clustering methods aim to segment spatial point datasets into various spatial cliques, which have been adopted by some studies to extract multi-scale accident-prone black spots [17,22,23,26,39]. Traditional clustering methods mainly contain partition clustering methods, density-clustering methods, hierarchical clustering methods, etc. [40]. However, the cluster number of partition clustering methods and the splitting threshold of hierarchical clustering methods can hardly be predetermined according to historical accident datasets. Differently, as a classical density-clustering method, DBSCAN first detects high-density core points and then iteratively aggregates the core points and their neighborhood points into individual clusters by the density-connectivity rules [41].

Compared to other clustering algorithms, DBSCAN is independent of the cluster numbers and the cluster sizes, which have been successfully used in some of the literature [26,42,43] and indicates some accuracy improvement of accident black spot identification. However, previous DBSCAN-based methods do not consider temporal factors and accident severities. Hence, this paper proposes a Spatiotemporal Density-based Accident Clustering (STDAC) method to identify varying-scale accident-prone black spots.

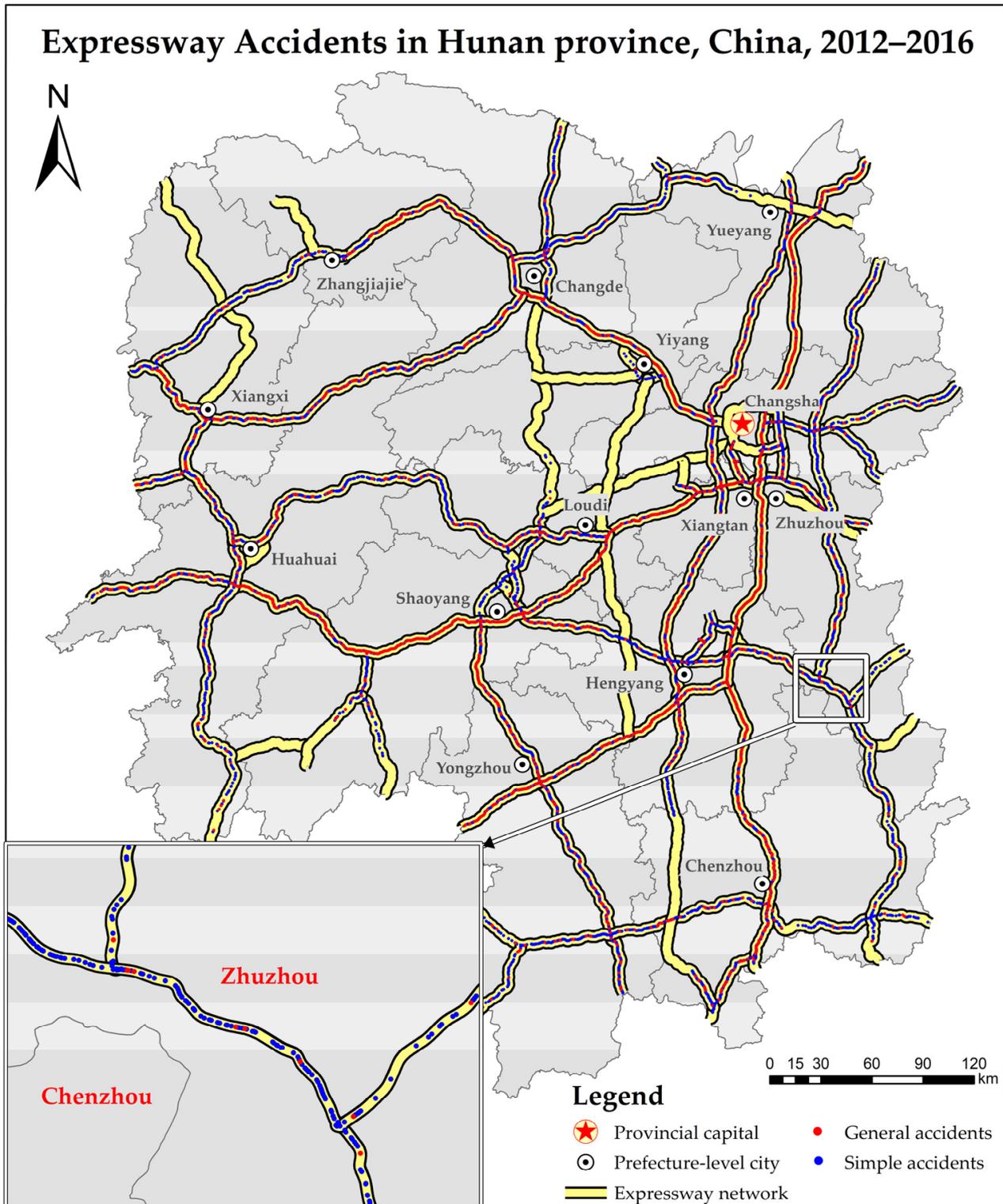
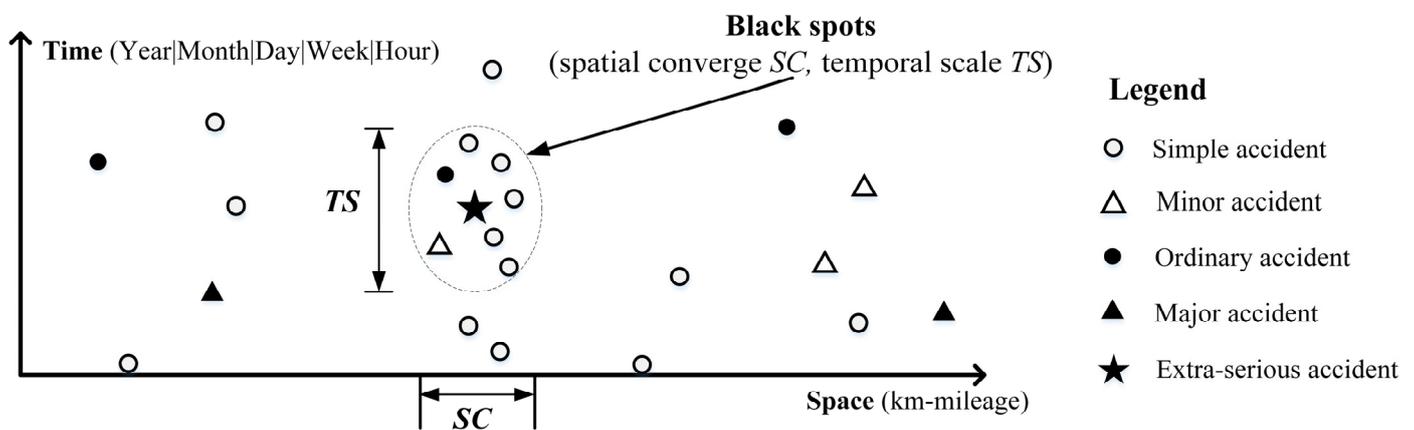


Figure 1. Geo-referenced accident data.

Various expressway black spots probably may show different spatiotemporal patterns of traffic accidents. For some black spots, the traffic accidents may concentrate on some fixed time period because of heavy weather, whereas the traffic accidents of some other black spots may show homogeneous temporal distribution on account of multiple causal features. To fully explore the spatiotemporal patterns of traffic accidents, the paper introduces the time periodicity into the definition of spatiotemporal neighborhood accidents and meanwhile considers the accident severity to count the neighborhood accidents for each accident. Based on the definition of spatiotemporal neighborhood accidents, the core accident points are detected and successively extended to discover the accident clusters of black spots according to classical density-connectivity rules. It can be seen from Figure 2 that the identified black spots based on STDAC contain two elements. One element is Spatial Coverage *SC*, which depicts the mileage range of the black spot. The other one is Temporal Scale *TS*, which indicates the accident-prone periods of black spots. The combination of spatial and temporal descriptors about black spots can provide important basis for making temporally targeted risk precaution solutions.



**Figure 2.** Spatiotemporal density-based accident clustering for black spot identification.

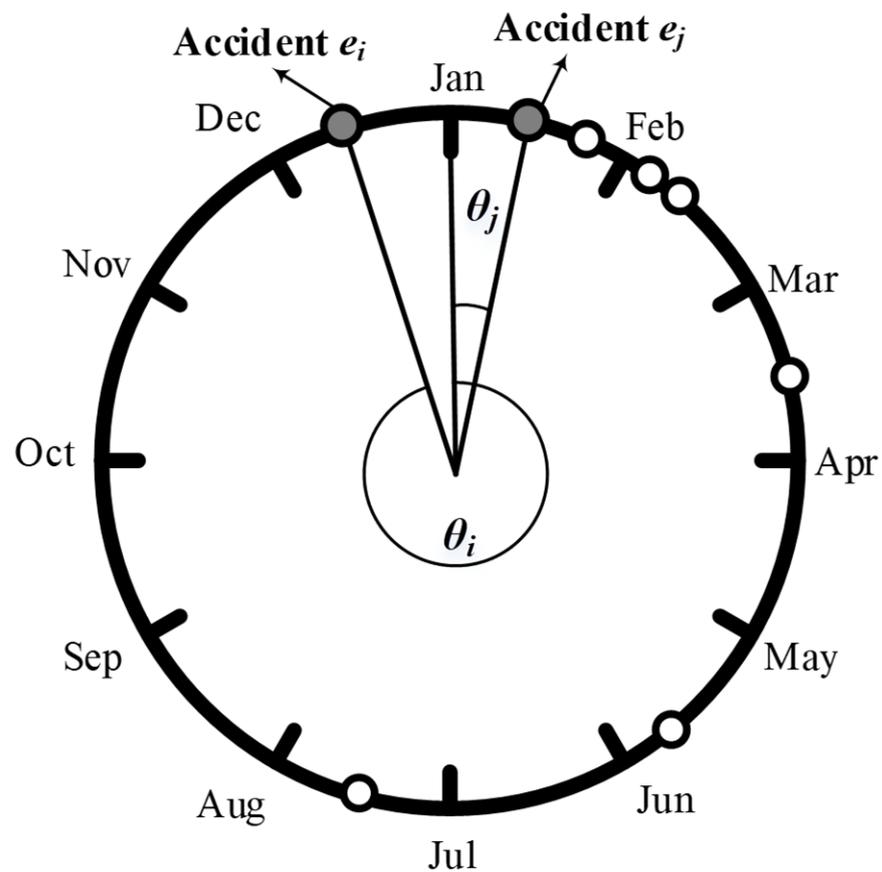
### 3.2.1. Definition of Spatiotemporal Neighborhood Accidents

The time periodicity is very common in our actual life, such as 23:00 PM being temporally close to 0:00 AM, December being close to January, etc. Similarly, many geographical phenomena also show obvious time periodicities, such as day and night, four seasons, daily commuting, etc. Hence, it is necessary to model the time periodicity of traffic accidents and further explore the potential time-related accident causes. In this paper, to mine the time periodicity of traffic accidents, the time components (e.g., month, day, week number, hour, minute, second) of each accident are separately extracted and then we convert accident time into an angle representation based on a predefined dial.

For example, as illustrated in Figure 3, 12 months are represented as an equally spaced dial, where January corresponds to  $0^\circ$ , December corresponds to  $330^\circ$ , and so on. For one accident  $e$ , the angle representation  $\theta$  is calculated as follows.

$$\theta = \left[ (T_M - 1) + \frac{T_d \times \varepsilon_d + T_h \times \varepsilon_h + T_m \times \varepsilon_m + T_s}{D \times \varepsilon_d} \right] \times 30^\circ \quad (1)$$

where  $T_M$ ,  $T_d$ ,  $T_h$ ,  $T_m$ , and  $T_s$  are the month, day, hour, minute, and second components of accident  $e$ , respectively.  $\varepsilon_d$ ,  $\varepsilon_h$ , and  $\varepsilon_m$  are the number of seconds in one day, one hour, and one minute, respectively (for instance,  $\varepsilon_m$  equals 60 s and  $\varepsilon_h$  equals 3600 s).  $D$  denotes the total number of days in the month of accident  $e$ , which possibly equals 28, 29, 30, or 31.



**Figure 3.** Angle partitioning of month component of traffic accidents.

As illustrated in Figure 3, the accident time of two accidents  $e_i$  and  $e_j$  are “13 December 2014 15:20:00” and “10 January 2012 12:10:00”, respectively. The angle values of  $e_i$  and  $e_j$  are calculated as  $\theta_i$  and  $\theta_j$ , which are represented as gray-filled circles in Figure 3. It is clearly seen that the angle representations of  $e_i$  and  $e_j$  are very approximate on the circular dial. Based on the angle representation of accident time, we can define the temporal distance  $Dis\_T$  between two accidents  $e_i$  and  $e_j$ , according to Formula (2) whereas the spatial distance  $Dis\_S$  is calculated as the Euclidean distance between two accident points.

$$Dis\_T(e_i, e_j) = \sin \frac{|\theta_i - \theta_j|}{2} \quad (2)$$

### 3.2.2. Counting of Spatiotemporal Neighborhood Accidents

Traditional DBSCAN needs to set two parameters of  $Eps$  and  $Minpts$ . Thereinto,  $Eps$  is a distance threshold for searching the neighborhood points.  $Minpts$  is a point number threshold for detecting the core points. In this paper, both spatial and temporal distance thresholds  $Eps_s$  and  $Eps_t$  are required to define the neighborhood accidents  $Neighbor(e_i)$  for each accident  $e_i$  by Formula (3).

$$Neighbor(e_i) = \{e_j | Dis\_S(e_i, e_j) \leq Eps_s \text{ and } Dis\_T(e_i, e_j) \leq Eps_t\} \quad (3)$$

In addition, DBSCAN simply selects these points with the number of neighborhood points larger than  $Minpts$  as core points. However, as different traffic accidents may lead to varying degrees of casualties and damage, the accident severity level should be integrated with the weights of the neighborhood accidents. With reference to the Road Traffic Accident Grade Standard issued by the Ministry of Public Security [44], we divide all traffic accidents

into five grades: 1st simple accident, 2nd minor accident, 3rd ordinary accident, 4th major accident, and 5th extra-serious accident according to the following rules.

- **1st Simple accident:** Accidents without injuries (i.e., simple accident dataset) are marked as simple accidents, of which the accident weight is uniformly assigned to  $w_1 = 1/3$ ;
- **2nd Minor accident:** Accidents with 1–2 slight injuries are marked as minor accidents, of which the accident weight is uniformly assigned to  $w_2 = 1$ ;
- **3rd Ordinary accident:** Accidents with 1–2 serious injuries or more than 3 slight injuries are marked as general accidents, of which the accident weight is uniformly assigned to  $w_3 = 3$ ;
- **4th Major accident:** Accidents with 1–2 deaths or 3–10 serious injuries are marked as major accidents, of which the accident weight is uniformly assigned to  $w_4 = 5$ ;
- **5th Extra-serious accident:** Accidents with more than 3 deaths, or more than 11 serious injuries, or 1 death and meanwhile more than 8 serious injuries, or more than 2 deaths and meanwhile more than 5 serious injuries are marked as Extra-serious accidents, of which the accident weight is uniformly assigned to  $w_5 = 7$ .

The weight values of minor accidents and above are set as an arithmetic progression according to the injuries or deaths, and the weight value of simple accidents is uniformly set as  $1/3$ . Based on the definition of accident severity and corresponding weights, the cardinality of neighborhood accident set for each accident  $e_i$  is calculated as  $|Neighbor(e_i)| = \sum_j N_j \times w_j$ , where  $N_j$  is the total number of  $i$ th-grade accidents in  $Neighbor(e_i)$  and  $w_j$  is the corresponding accident weight for  $i$ th-grade accident, respectively.

### 3.2.3. Identification of Accident-Prone Black Spots Based on DBSCAN

Through introducing the temporal periodicity and improving the counting of neighborhood accidents, the Spatiotemporal Density-based Accident Clustering process based on DBSCAN is implemented as follows:

Step 1: Detect the core accidents by judging whether the cardinalities of their neighborhood accidents are larger than  $Minpts$ ;

Step 2: Randomly select an unprocessed core accident and add the core accident together with its neighborhood accidents into a newly created cluster;

Step 3: Recursively traverse all the unprocessed core points in the current cluster and insert the density-reachable accidents of the traversed accident into the current cluster until all core accidents in the current cluster are processed;

Step 4: Return to Step 2 until all the core points have been processed.

After spatiotemporal density-based accident clustering (STDAC), the weighted accident count and temporal accident frequency of each cluster is calculated and ranked for identifying accident-prone black spots.

### 3.3. Divide-and-Conquer Accident Causal Analysis Based on Bayesian Network

As mentioned before, a traffic accident is an unlikely event caused by many factors. Bayesian Network (BN) utilized a directed acyclic graph (DAG) model to represent the conditional correlations between variables, providing a feasible tool for uncertainty reasoning and data analysis. In the Bayesian Network, each node represents an attribute variable and each directed edge corresponds to the conditional dependency between the two connected nodes. Previous studies [36,37,45] have demonstrated that BN is applicable to exploring the potential relationship between accident causal factors. However, the existing research applied all the observed accidents to construct a uniform Bayesian Network, which can hardly profile the spatial heterogeneities of accident causal factors in different regions. Hence, the paper proposes to adopt a divide-and-conquer strategy that is separately constructing different Bayesian networks for accident-prone black spots and the whole study area to explore their accident causal patterns. In general, a Bayesian Network  $B$  contains two parts, namely structure  $G$  and parameter  $\Theta$ . Hence, The BN-based accident causal inference process consists of two key steps, i.e., structure learning and parameter learning.

Structure learning is to determine the topological graph structure  $G$  of accident Bayesian Network. Parameter learning is to acquire the quantitative parameters to describe the conditional probability table (CPT) between two connected nodes.

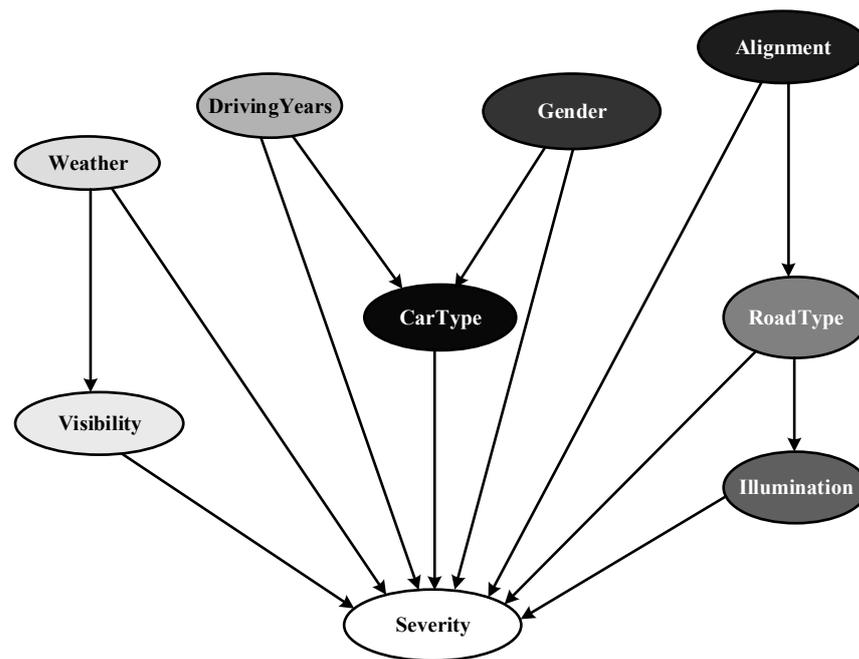
### 3.3.1. Structure Learning of Accident Bayesian Network

Generally, there are expert-based methods, machine-learning methods, and hybrid methods to determine the directed acyclic graph structure of Bayesian Network [36]. The expert-based method directly establishes Bayesian Network by absorbing experts' opinions and the machine-learning method automatically learns Bayesian Network structure from observed data. For balancing the dependence on experts' knowledge and the interpretability of auto-learned models, we incorporate the TAN-based learning structure [46] and experts' knowledge to construct accident Bayesian Network. As listed in Table 2, 9 accident nodes are selected to constitute accident Bayesian Network to be built. Thereinto, the node of accident severity is set as target node in order to analyze other factors' impact on accident severity. Weather, Visibility, Illumination, and Alignment are directly obtained from general accident data. Car Type, Gender, and Driving Years are generated from the personal information of the participant with main accident liability. Particularly, Road Type is determined as either Road Sections or Road Facilities by judging whether the accident point is located within the buffer region of some road facilities, such as junctions, service zones, and tollgates. Finally, severity is calculated according to the schemes of Section 3.2.2.

**Table 2.** The nodes constitute Bayesian Network and the corresponding variable values.

No.	Node	Variable
1	CarType	Small Passenger Car; Middle Passenger Car; Large Passenger Car; Light Truck; Medium Truck; Heavy Truck; Pedestrian; Other
2	Weather	Sunny; Rainy; Cloudy; Other
3	RoadType	Road Sections; Road Facilities (e.g., junctions, service zones, tollgates)
4	Gender	Male; Female
5	DrivingYears	type1: $\leq 1$ year; type2: 2–5 years; type3: 6–10 years; type4: $> 10$ years; no-licensed
6	Visibility	type1: $\leq 50$ m; type2: 50–100 m; type3: 100–200 m; type4: $> 200$ m
7	Illumination	Day; Dusk; Dawn; No Lighting at Night; Lighting at Night
8	Alignment	Flat; Bend; Slope; Sharp Bend; Steep Slope; Bend and Slope; Sharp Bend and Slope; Steep Slope and Bend; Sharp Bend and Steep Slope; Long Descent
9	Severity	Simple Accident; Minor Accident; Ordinary Accident; Major Accident; Extra-serious Accident

The accident Bayesian Network is constructed and shown in Figure 4. The prior probability of each node variable is calculated by counting the corresponding ratio of general accidents. As shown in Figure 4, the arrowed line indicates one causal relation exists in the connected nodes. The darker color of nodes denotes it has more important impact on the target nodes of accident severity. It can be seen that the top four factors impacting accident severity are Car Type, Alignment, Gender, and Illumination, respectively. They are the representative factors concerned with vehicle-related, road-related, driver-related, and environment-related features, respectively.



**Figure 4.** Result of structure learning of accident Bayesian Network.

### 3.3.2. Parameter Learning of Accident Bayesian Network

After structure learning of accident Bayesian Network, parameter learning is performed to quantitatively obtain the parameters for describing accident Bayesian Network. Suppose the learned directed acyclic graph structure is  $G = \{V, E\}$ ,  $V$  contains  $d$  nodes, i.e.,  $X_1 \dots X_d$ , the possible values for each node  $X_i$  are in the set  $\chi_i = \{x_i^1 \dots x_i^{k_i}\}$ , and the parent configuration  $\pi_i$  for each node  $X_i$  is in the set  $\Pi_i = \{\pi_i^{(1)} \dots \pi_i^{(q_i)}\}$ , where  $q_i = \prod_{X_h \in \text{parent}(X_i)} k_h$ . The conditional probability  $\theta_{ji,l}$  to be estimated is defined as follows:

$$\theta_{ji,l} = P(X_i = x_i^j | \pi_i = \pi_i^{(l)}), \quad i = 1 \dots d, \quad j = 1 \dots k_i, \quad l = 1 \dots q_i \quad (4)$$

It is obvious to meet  $\sum_{j=1}^{k_i} \theta_{ji,l} = 1$ . The paper utilized maximum likelihood estimation (MLE) method to determine the parameter variable  $\theta_{ji,l}$  according to Formula (4).

$$\delta_{ji,l}^{MLE} = \frac{N_{ijl}}{N_{il}}, \quad N_{il} = \sum_{j=1}^{k_i} N_{ijl} \quad (5)$$

where  $N_{ijl}$  denotes the number of general accident data that meets both  $X_i = x_i^j$  and  $\pi_i = \pi_i^{(l)}$ .

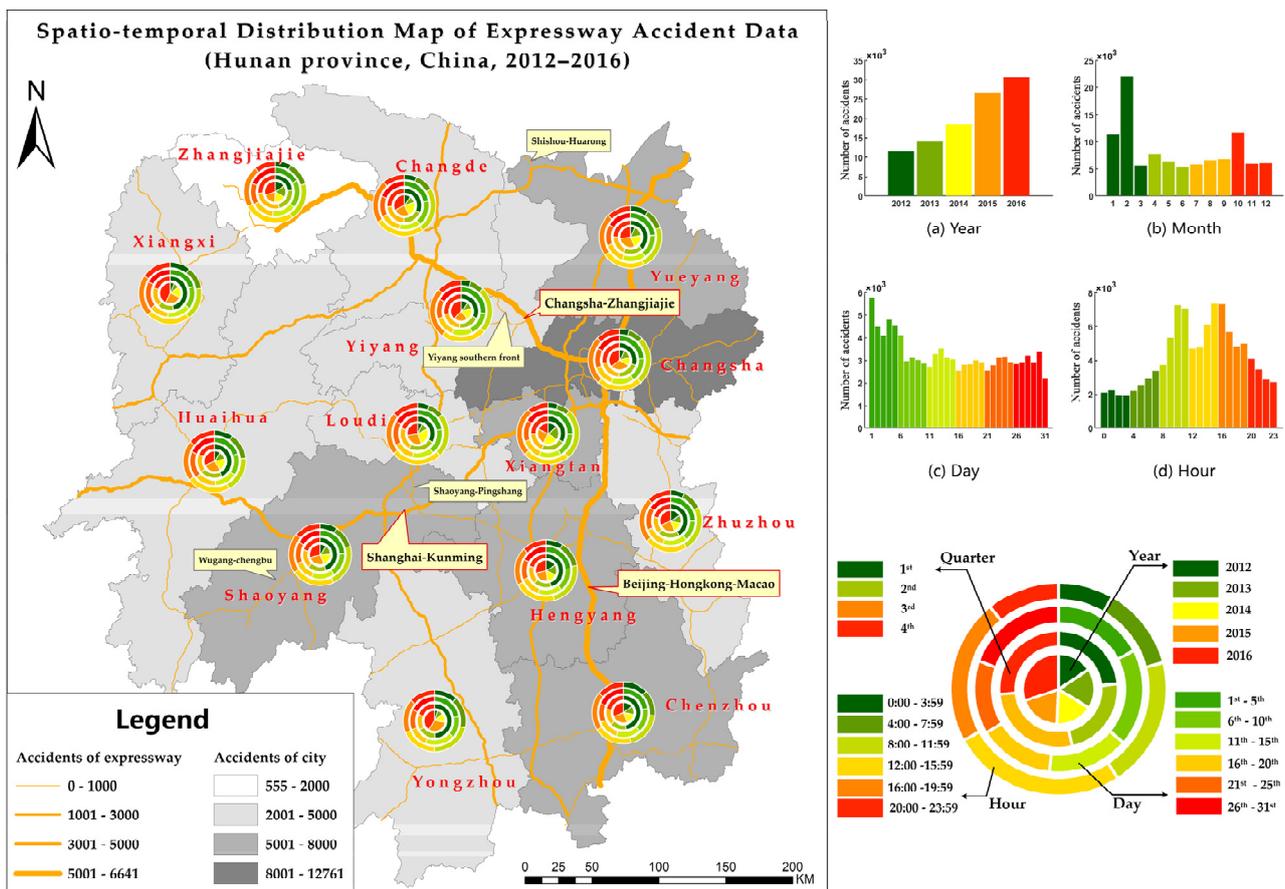
Based on Bayesian Network, reverse inference is used to explore the potential causal factors for different black spots and the whole expressway network. Theoretically, reverse inference is used to calculate the posterior probability of each parent node variable under a target node variable and then compare the posterior probability with its prior probability, in order to find the potential causal factors leading to the target node variable [47].

## 4. Results

In the experimental analysis, an exploratory analysis of expressway accident data is first conducted by overlapping the geo-referenced accident data with the Hunan expressway network and the administrative areas of prefecture-level cities. After that, the results of identifying accident-prone black spots and exploring accident causal factors are qualitatively and quantitatively analyzed to verify our proposed method.

#### 4.1. Exploratory Analysis of Expressway Accident Data

Figure 5 shows a spatiotemporal distribution map of expressway accident data in Hunan province, China, during 2012–2016. Firstly, we find that there are obviously spatial aggregation patterns of expressway accidents both at the expressway-level scale and prefecture-level city scale. At the prefecture-level city scale, more accidents happened in the eastern cities and relatively fewer accidents happened in the western cities. Specifically, as the capital of Hunan province, Changsha has the most accidents, whereas the city with the fewest accidents is Zhangjiajie in the west, which is one of the most famous tourist cities in Hunan. At the expressway level, the Beijing-Hongkong-Macao expressway (G4), Shanghai-Kunming expressway (G60), and Changsha-Zhangjiajie expressway (G5513) are the top three expressways (see the text annotations with red borders) with highest accident volume. The representative expressways with the fewest accidents (see the text annotations with black border) are Yiyang southern front, Shishou-Huarong expressway, Shaoyang-Pingshang expressway, and Wugang-chengbu expressway, which are mostly provincial-level expressways, and about 40 accidents took place in the observed five years. Overall, the spatial patterns illustrated in historical expressway accident data may be related to the spatial variations in the social economy, human traveling, expressway density, and traffic flow between different regions.



**Figure 5.** Spatiotemporal distribution map of expressway accident data in Hunan Province during the period of 2012–2016.

Secondly, statistical analysis of expressway accidents at four temporal granularities (i.e., yearly, monthly, daily, and hourly) are elaborated in the upper right histograms of Figure 5, respectively. At the yearly granularity, the annual expressway accidents have increased year after year, and the number of accidents that occurred in 2016 was double the number of accidents that occurred in 2012. Large challenges for traffic safety management

have been caused by the popularization of motor vehicles and the increasing acceleration of urbanization. At the monthly granularity, there are two obvious peaks of expressway accidents (i.e., Jan-Feb and Oct), that account for about 44.28% of the total accidents. It is found that these two peak periods correspondingly cover the two biggest holidays in China (i.e., Spring Festival and National Day) when larger human traveling occurs and raises the accident risk. At the daily granularity, the daily accidents in the month are almost uniformly distributed, with slightly more accidents at the beginning days of one month. Finally, at the hourly granularity, there are also two obvious peaks of expressway accidents (i.e., 10:00–12:00 and 16:00–17:00), which account for about 48.85% of all accidents and are possibly relevant to the human daily schedule.

Thirdly, we make a spatiotemporal integration analysis for each prefecture-level city. In detail, we compute the yearly, quarterly, daily, and hourly expressway accident proportions for each prefecture-level city and intuitively plot the proportions as a circular pie chart located at the center of each prefecture-level city. As shown in Figure 5, although the temporal distributions of different prefecture-level cities were similar to the whole of the Hunan province, there were still some particular characteristics for individual prefecture-level cities, summarized as follows.

- (1) At the yearly granularity (shown as the innermost layer of the pie chart), some prefecture-level cities (Zhangjiajie, Hengyang, Zhuzhou, etc.) had relatively similar accident volumes in the observed 5 years but some other prefecture-level cities (Changsha, Loudi, Yongzhou, etc.) manifested a significant accident increase in the latter 3 years.
- (2) At the quarterly granularity (shown in the second layer of the pie chart), the quarterly expressway accident proportions in most prefecture-level cities were roughly similar but the expressway accidents of some special prefecture-level cities (e.g., Xiangxi, Huaihua, Yongzhou, and Chenzhou) mainly concentrated on the first quarter (nearly account for 50%).
- (3) At the daily granularity (shown in the third layer of the pie chart), the daily expressway accidents of most prefecture-level cities were nearly similar to each other, which also coincided with the daily distribution of Hunan province.
- (4) At the hourly granularity (shown in the outermost layer of the pie chart), the hourly expressway accidents were conformably displayed as daytime-dominant distributions for most prefecture-level cities.

#### 4.2. Results Analysis of Identifying Accident-Prone Black Spots

Based on the clustering parameters,  $Eps_s = 700$  m,  $Eps_t = 0.025$ , and  $Minpts = 10$ , the expressway accident-prone black spots are identified and depicted in Figure 6. In Figure 6, the identified accident-prone black spots as denoted as gray circles, of which the central point indicates the central accident position of each black spot and the circle size measures the weighted accident number of each black spot during the period of 2012–2016. The mileage range of each black spot and the belonging expressway route number are labeled in the nearby rectangle annotation, respectively. Meanwhile, different fill colors of the rectangle annotations are used to distinguish temporally clustering ranges of black spots. For example, the red color indicates that the black spot mainly had accidents in February and a combination of red and blue means the black spot mainly had accidents in February and October.

It can be seen from Figure 6 that different spatiotemporal patterns indeed existed in the identified black spots. At first, spatially, the identified black spots were mainly located at the Beijing-Hongkong-Macao expressway (G4), Shanghai-Kunming expressway (G60), and Changsha-Zhangjiajie expressway (G5513). The identified black spots on the above three expressways account for about 78.13% of all identified black spots, showing obvious spatial aggregation patterns of expressway black spots. Then, temporally, the accidents happening on various black spots were intensively at different time periods,

showing significant temporal heterogeneities of accident-prone periods between different black spots.

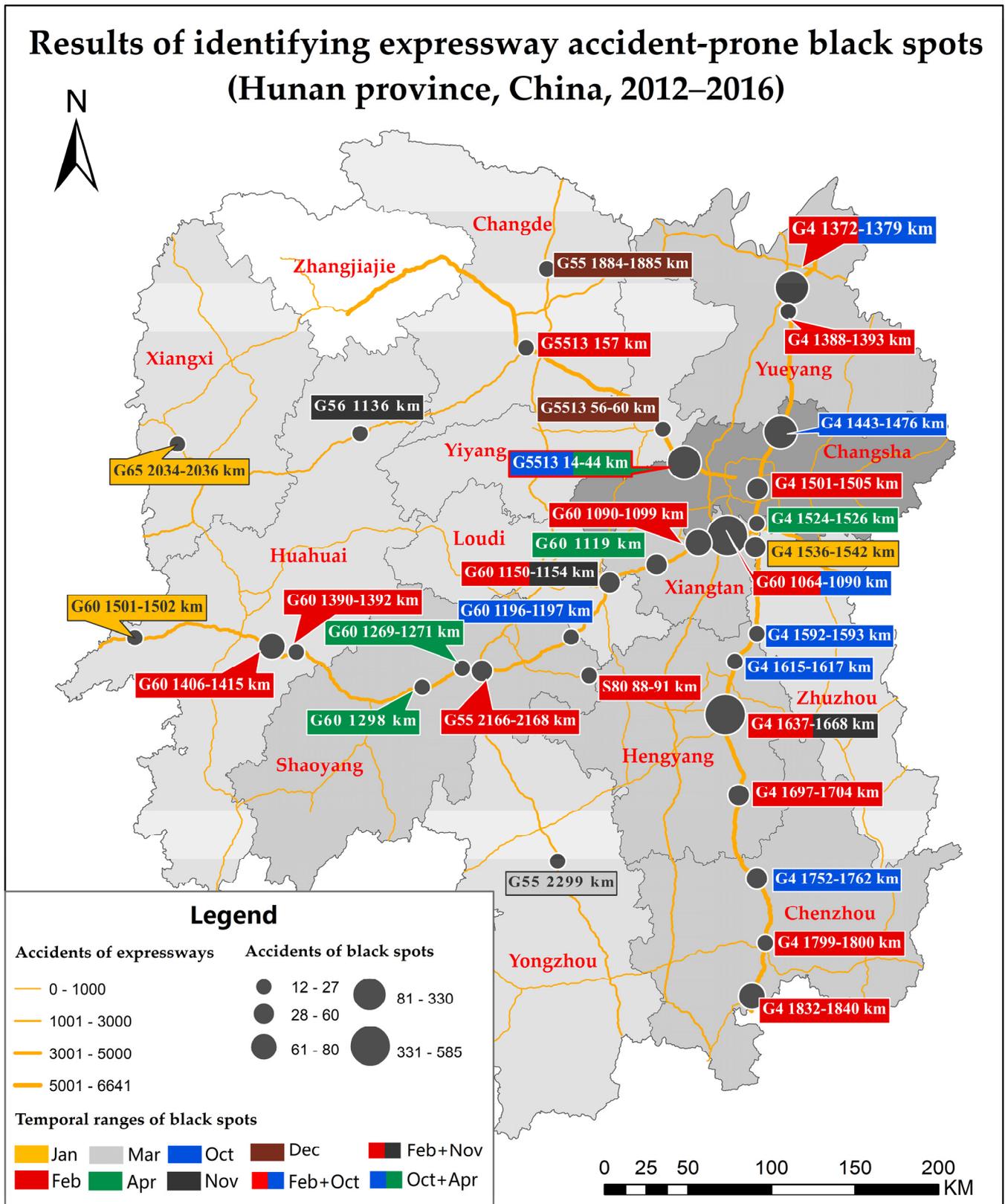


Figure 6. Results of identifying expressway accident-prone black spots in Hunan province, China.

Table 3 lists the mileage range, weighted accident volume, and accident proportion of typical months for the top four accident-prone black spots identified by the proposed method. As listed in Table 3, for the top black spot (G60 1064–1090 km), about 63.77% of accidents were happening in February (32.69%) and October (31.08%). However, for another two black spots, G4 1637–1668 km and G4 1443–1476 km, the accidents mostly happen in February (68.22%) and October (87.65%), respectively. It demonstrates that the proposed method can adaptively identify accident-prone black spots with different spatial coverages and temporal ranges, efficiently dealing with the problem of MAUP in black spot identification. Moreover, the temporal periodicity and accident severity are introduced into spatiotemporal density-based accident clustering, which can effectively make up for the confronting deficiency of identifying spatiotemporal black spots. Various spatiotemporal patterns of black spots may put forward urgent demands for divide-and-conquer accident causal analysis and targeted traffic security treatments in future research.

**Table 3.** Accident volume and proportion statistics for the top four expressway black spots.

Mileage Range of Accident-Prone Black Spots		Weighted Accident Count	Accident Proportion of Typical Month(s)
Shanghai-kunming (G60)	1064–1090 km	558	Feb (32.69%) + Oct (31.08%) = 63.77%
Beijing-Hongkong-Macao (G4)	1637–1668 km	335	Feb: 68.22%
Changsha-Zhangjiajie (G5513)	14–44 km	283.7	Oct: 59.97%
Beijing-Hongkong-Macao (G4)	1443–1476 km	193.3	Oct: 87.68%

According to one official report of 2017 by the traffic management department [48], the Changsha-Zhangjiajie expressway 15–33 km (G5513 15–33 km) is recognized as one of the top 10 accident-prone black spots in China, which is demonstrated to be very consistent with the black spot of G5513 14–44 km identified by our method. Further analysis in Figure 7 reveals that there are as many as 10 road facilities along this black spot in spite of only 30 km mileage. Besides passing through many road facilities (e.g., bridges, tollgates, service areas, and junctions), the black spot is also particularly located at the exit and entrance roads between two important cities of Changsha and Ningxiang, leading to its complex road condition and heavy traffic flow.

Additionally, Figure 8 illustrates the statistical histogram of traffic accidents on the identified black spots and other road sections of G4 expressway in Hunan province (5 km per bin). The red bins above the horizontal axis denote the black spots identified by our method and the yellow bins below the horizontal axis denote the black spots released by the official department [49]. It can be seen that seven of nine black spots published to the public are correctly identified by our method and have common sections as long as 55 km. Relatively high consistency between our results and the official results demonstrates the good performance of our density-clustering method.

To contrastively analyze the accuracy of black spot identification, we apply the proposed method to two accident datasets concerning the Beijing-Hongkong-Macao expressway (G4) during 2012–2016 and in the year of 2018. As shown in Figure 9, the results of black spot identification based on two accident datasets have as long as 75 km of common segment parts, indicating high spatial consistency between the identified black spots. Additionally, we also find that the temporal ranges of black spots identified by two accident datasets are similarly concentrating on two periods, i.e., mid-January to early March and early September to early October. Particularly, there are two black spots *i* and *j* identified by the 2012–2016 accident dataset (see the blue rectangles of Figure 9a), which have disappeared as a result of black spot identification in 2018 (see Figure 9b). Actually, Figure 10 shows that black spot *i* is located at the upstream road segments near the Leiyang service area in Hengyang, while black spot *j* is located at the road interchange of Zhuting tollgate.

It is valuable for other similar black spots to confirm whether efficient measures have been taken on these black spots.



Figure 7. The satellite imagery and road facilities along black spot Changsha-Zhangjiajie expressway G5513 14-44 km.

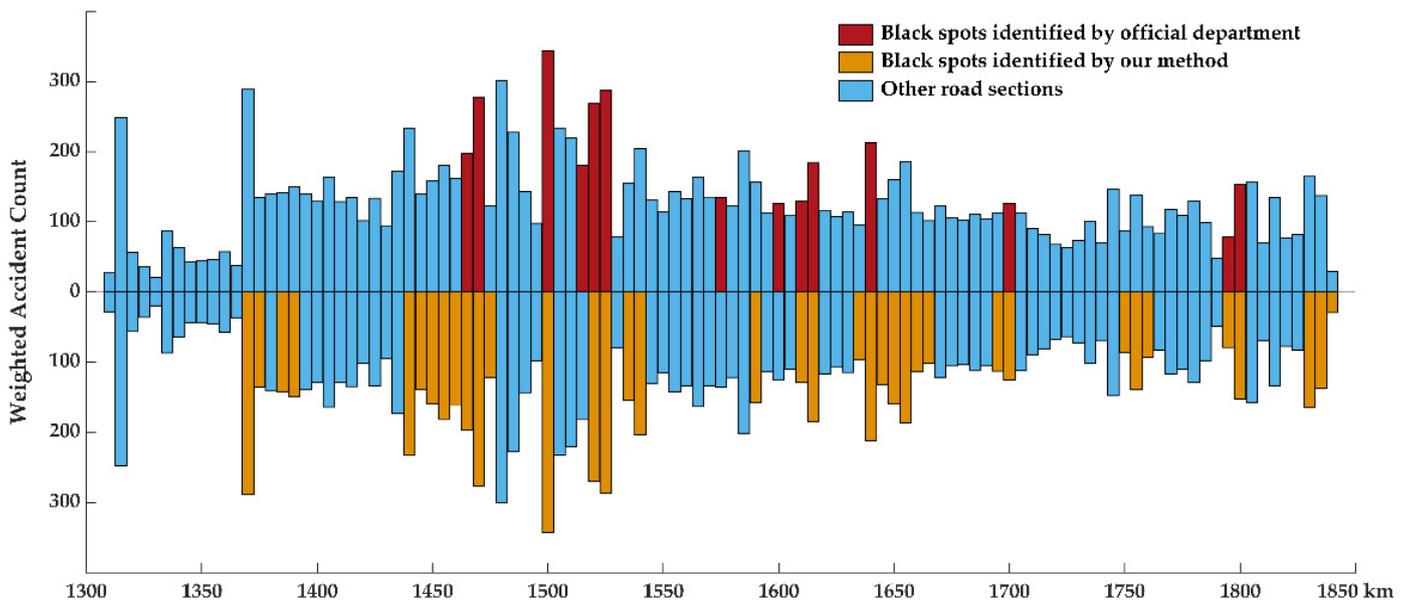
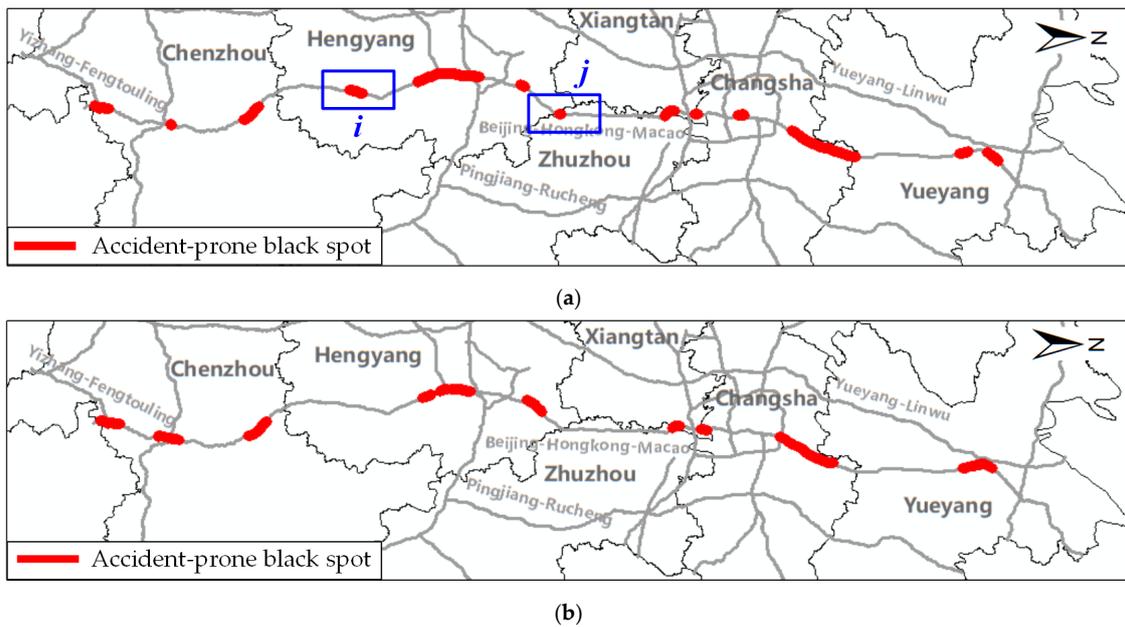
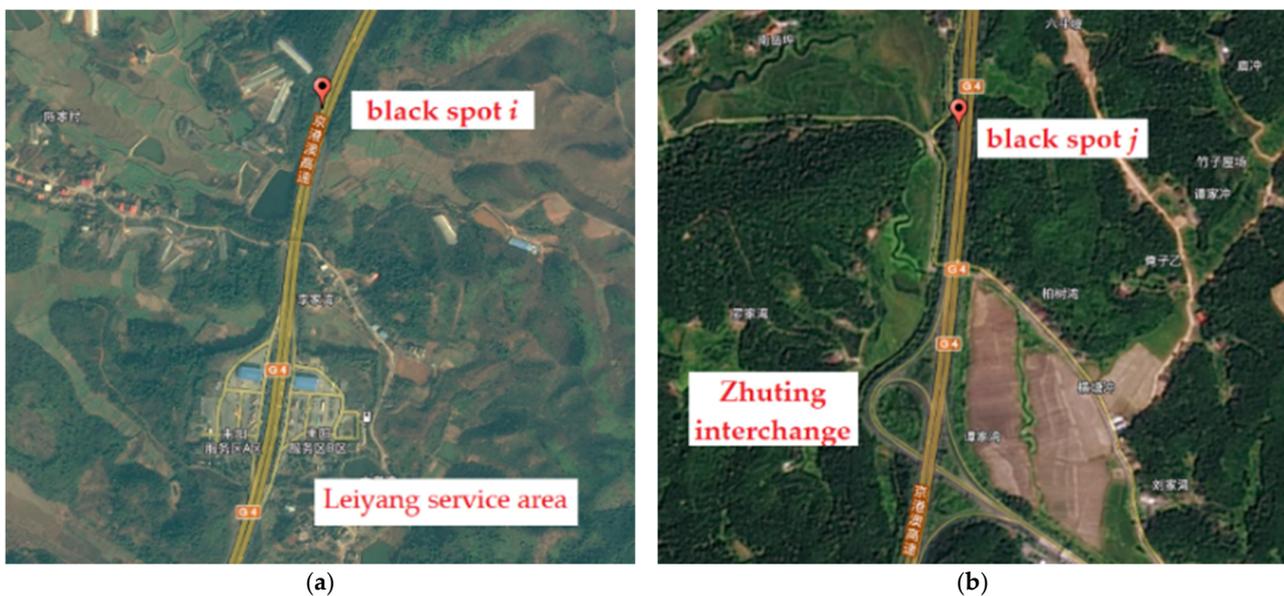


Figure 8. Statistical histogram of traffic accidents in Hunan sections of G4 expressway.



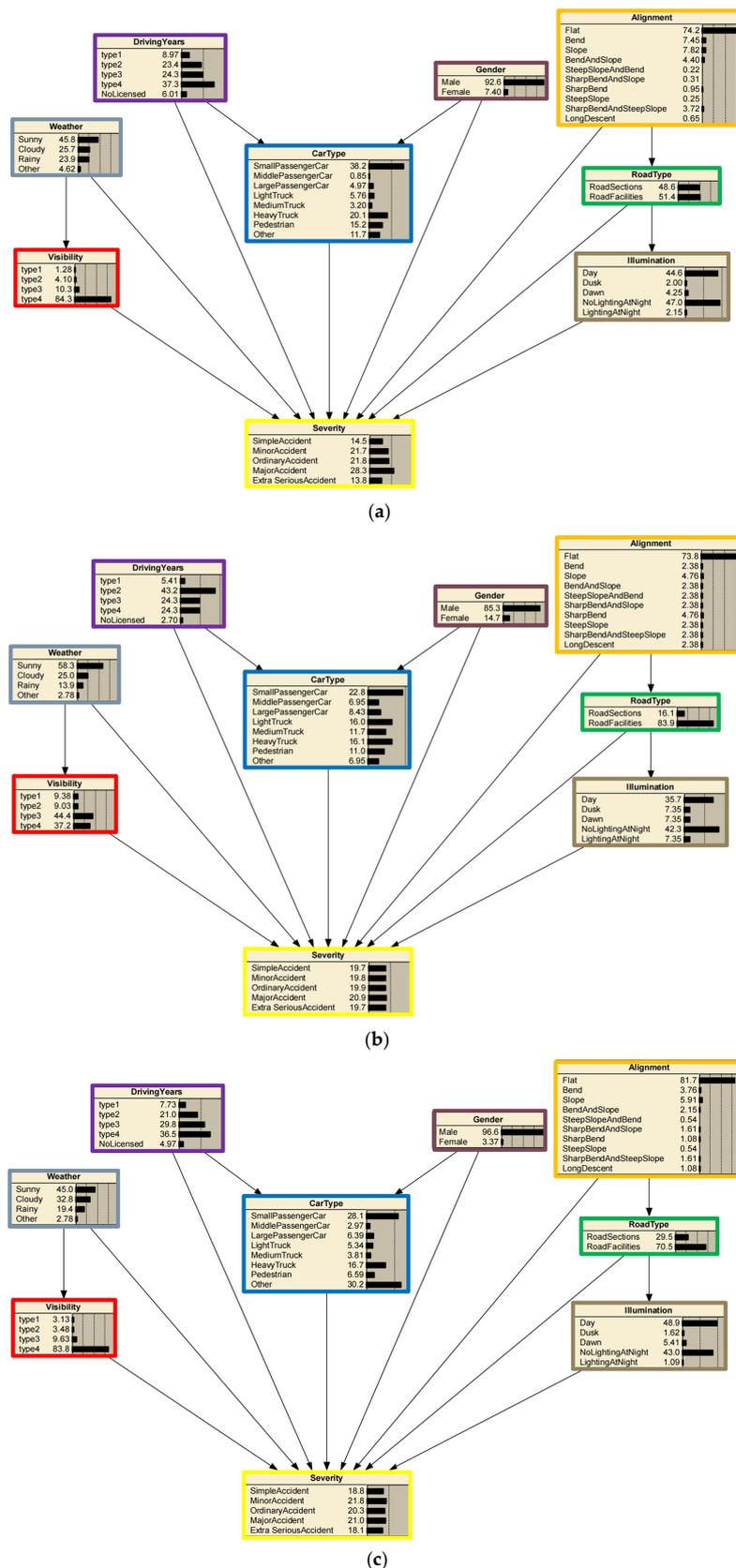
**Figure 9.** Results comparison of the identified black spots on Beijing-Hongkong-Macao expressway using the accident data of 2012–2016 and 2018. (a) The identified black spots of Beijing-Hongkong-Macao expressway using 2012–2016 dataset. (b) The identified black spots of Beijing-Hongkong-Macao expressway using 2018 dataset.



**Figure 10.** The satellite map and road facilities around black spots *i* and *j* in Figure 8a. (a) black spot *i*; (b) black spot *j*.

#### 4.3. Results Analysis of Exploring Accident-Prone Causal Factors

In this paper, we conduct reverse inference analysis for extra-serious accidents using Netica software and attempt to explore heterogeneous accident factors for various black spots and the whole expressway network. We select two black spots identified by our method, i.e., Changsha-Zhangjiajie expressway 14–44 km (G5513 14–44 km) and Beijing-Hong Kong-Macao expressway 1637–1668 km (G4 1637–1668 km), to compare their heterogeneous causal factors. Figure 11a–c depict the prior probabilities of all node variables for the whole Hunan expressway network, black spots G5513 14–44 km, and G4 1637–1668 km, respectively.



**Figure 11.** Prior probabilities of all node variables for various black spots and the whole area. (a) Prior probabilities of all node variables for the whole expressway network. (b) Prior probabilities of all node variables for black spot G5513 14–44 km. (c) Prior probabilities of all node variables for black spot G4 1637–1668 km.

At first, it can be roughly seen from Figure 11 that the prior probabilities of the identical node show some discrepancies between the whole expressway and black spots. On the one hand, it can be seen from the green node Road Type that 83.9% of accidents on black spot G5513 14–44 km and 70.5% of accidents on black spot G4 1637–1668 km were located at road facility sections while only 51.4% of accidents in the whole expressway network were located near road facilities. Various black spots and expressways show different spatial distributions of facility-related accidents, which demand the formulation of targeted road-designing solutions for traffic risk precaution. On the other hand, we find from the blue node Car Type that only some vehicle types (e.g., small passenger cars) are the main accident participants in the Hunan expressway network and black spot G4 1637–1668 km, whereas relatively more vehicle types are likely to happen road accidents on black spot G5513 14–44 km. It indicates that various black spots and expressways may have the main accident participants of their own. Especially, a lot of road facilities (see Figure 7) are distributed along black spot G5513 14–44 km, which probably leads to common accident proneness for most vehicle types. Moreover, according to the red node of visibility, we found from black spot G5513 14–44 km that its prior probabilities of low visibility (e.g., type1–3: <200 m) are larger than that of the whole area and black spot G4 1637–1668 km, indicating that bad visibility environment may exacerbate the traffic risk of G5513 14–44 km. In addition, compared with the whole Hunan expressway network, the number of female-related accidents on black spot G5513 14–44 km shows a significant increase but the corresponding accident number on blackspot G4 1637–1668 km decreased. Finally, we analyze the target node of accident severity and discover that the proportions of extra-serious accidents on black spots G5513 14–44 km and G4 1637–1668 km are relatively higher than that of the whole expressway network, indicating the potential accident risk on black spots.

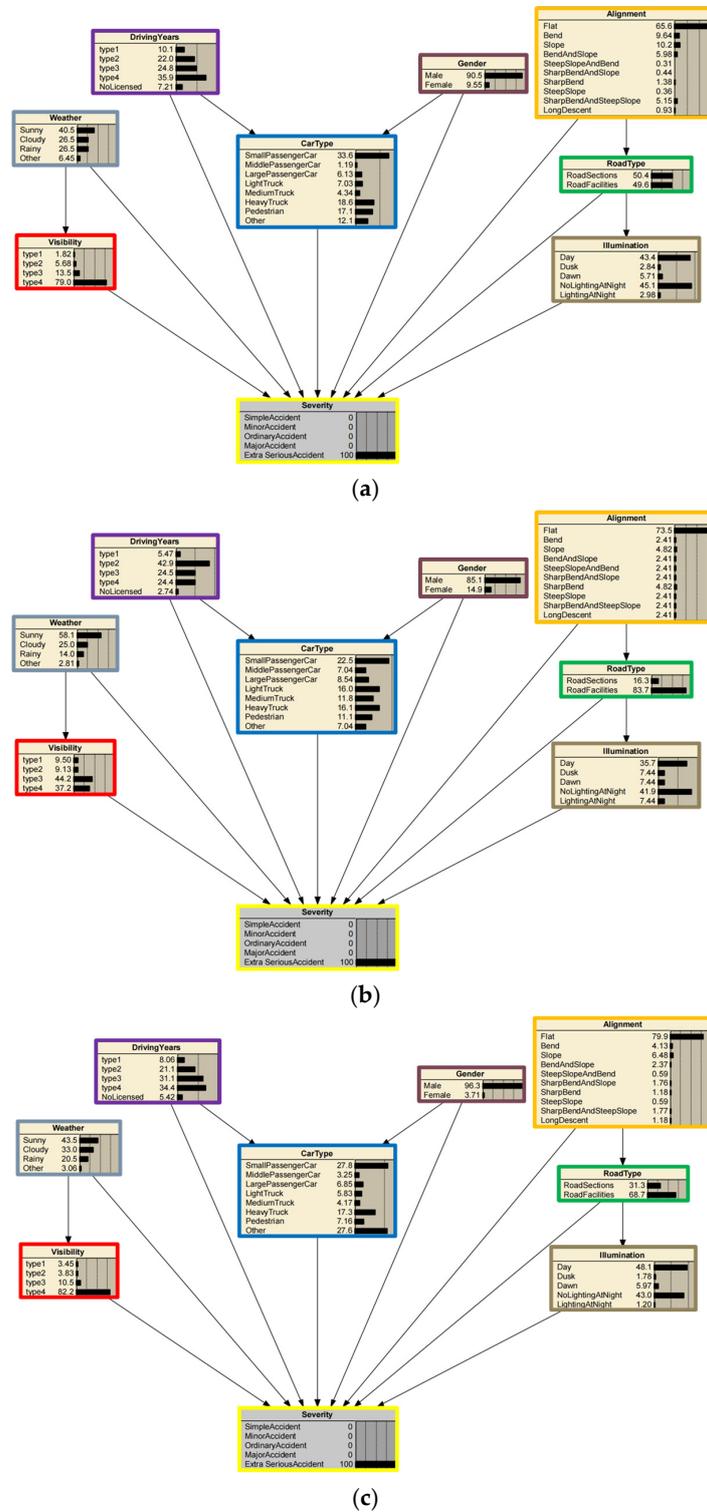
Figure 11b,c and Figure 12a contrastively illustrate the posterior probabilities of all node variables when extra-serious accidents happen in the whole Hunan expressway network, at black spots G5513 14–44 km and G4 1637–1668 km. Compared to Figures 11a–c and 12a–c, the identical node variables are showing minor differences between their prior and posterior probabilities. Hence, we calculate an increasing percentage  $IP_{i,j}$  by Formula (5) to analyze the causal impacts of different node variables on extra-serious accidents quantitatively.

$$IP_{i,j} = \frac{P(X_i = x_i^j | X_{severity} = x_{severity}^{Extra-serious accident}) - P(X_i = x_i^j)}{P(X_i = x_i^j)} \quad (6)$$

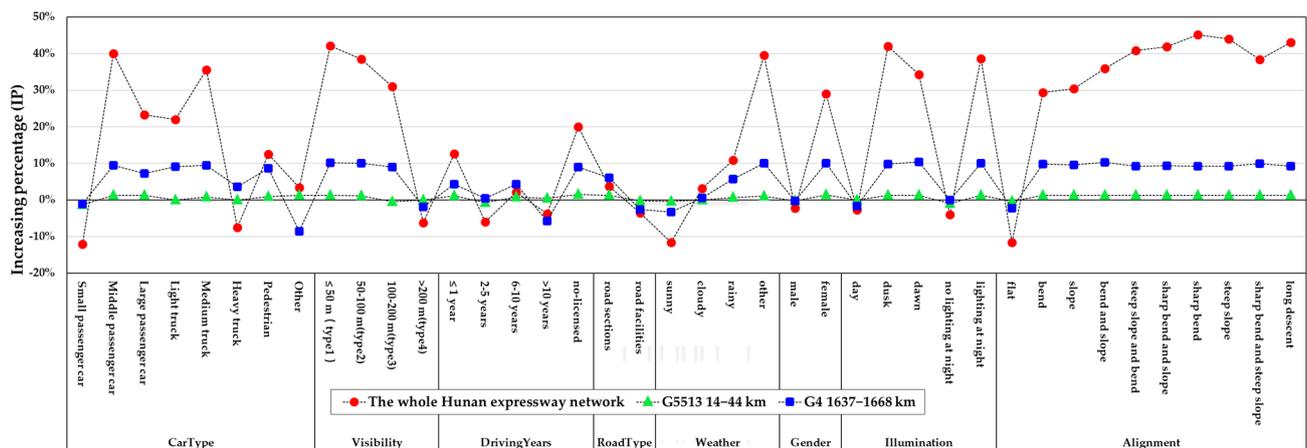
where  $P(X_i = x_i^j)$  means the prior probability that node  $X_i$  takes the state of  $x_i^j$ , and  $P(X_i = x_i^j | X_{severity} = x_{severity}^{Extra-serious accident})$  denotes the posterior probability that node  $X_i$  takes the state of  $x_i^j$  when extra-serious accidents happen (i.e.,  $X_{severity} = x_{severity}^{Extra-serious accident}$ ).

Figure 13 plots the increasing percentages ( $IP$ ) of different node variables when extra-serious accidents happen. The result of the whole expressway network is represented as red dotted lines. The results of two black spots G5513 14–44 km and G4 1637–1668 km are represented as green dotted lines and blue dotted lines, respectively. In general, the increasing percentages of G4 1637–1668 km are basically consistent with but lower than that of the whole expressway network. Contrastively, the increasing percentages of G5513 14–44 km keep a low level for almost all node variables, perhaps because of only a small number of general accidents recorded at this black spot. For the whole expressway network and G4 1637–1668 km, the posterior probabilities of some car types when extra-serious accidents happen, such as middle and larger passenger cars, light and medium trucks, and pedestrian (as well as heavy trucks on G4 1637–1668 km and other car types in the whole expressway network), have obviously increased more than their prior probabilities, indicating that the above-mentioned car types may increase the risk of extra-serious accidents. On the one hand, middle and larger passenger cars often carry more people than small passenger cars, which possibly leads to extra-serious accidents with more injuries

and heavy damages. On the other hand, the problems of overloading, long driving, and high speed are becoming more and more severe, especially for light and middle trucks, which amplify the risk and severity of extra-serious accidents.



**Figure 12.** Posterior probabilities when Extra-serious accidents happen in the whole area or at various black spots. (a) Posterior probabilities when Extra-serious accidents happen in the whole expressway network. (b) Posterior probabilities when Extra-serious accidents happen at black spot G513 14–44 km. (c) Posterior probabilities when Extra-serious accidents happen at black spot G4 1637–1668 km.



**Figure 13.** The increasing percentages (IP) of different node variables when Extra-serious accidents happen.

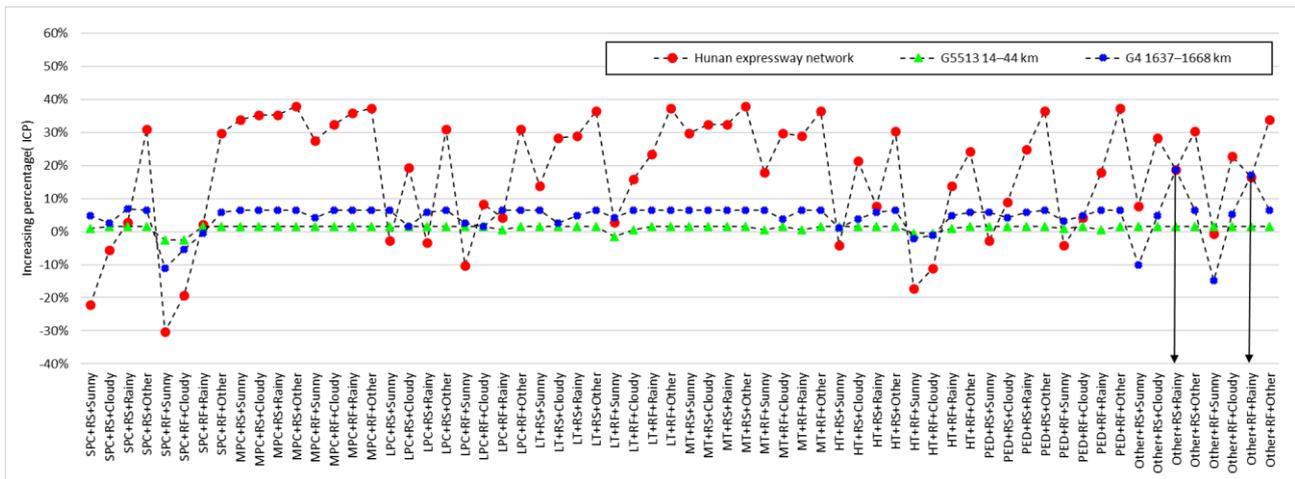
Similarly, we can see from Figure 13 that the posterior probabilities of visibility less than 200 m (visibility), no more than one driving year or even without a license (driving years), female drivers (Gender), rainy and other extreme weather (snowy, foggy, etc.), and weak illumination (e.g., dusk, dawn, and lighting at night) also show positive values of increasing percentage (IP) when extra-serious accidents happen, indicating more impacts of these factors on extra-serious accidents. Additionally, the non-flat road alignments (especially sharp bends, steep slopes, and long descents) show a high increasing percentage of posterior probabilities when extra-serious accidents happen. Those node variables with obviously increasing percentages should be paid more attention to accident risk prevention.

To quantitatively analyze the impacts of different conditional combinations on accident severity, we calculate the increasing percentage  $ICP_{i,j}$  between the conditional and prior probabilities as follows.

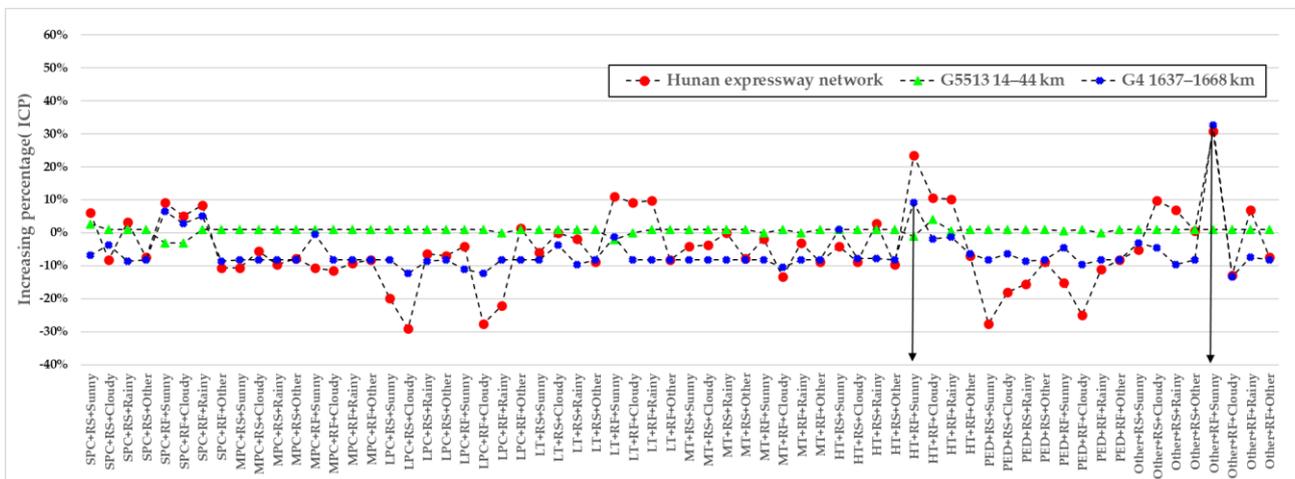
$$ICP_{i,j} = \frac{P(X = x^i | \Pi = \pi^j) - P(X = x^i)}{P(X = x^i)} \quad (7)$$

where  $P(X = x^i)$  is the prior probability of  $i^{th}$  accident severity and  $P(X = x^i | \Pi = \pi^j)$  denotes the conditional probability of  $i^{th}$  accident severity when  $\Pi$  takes the combination of  $\pi^j$ . In this paper, we analyze the combinations  $\Pi = (\text{Car Type}, \text{Road Type}, \text{and Weather})$ .

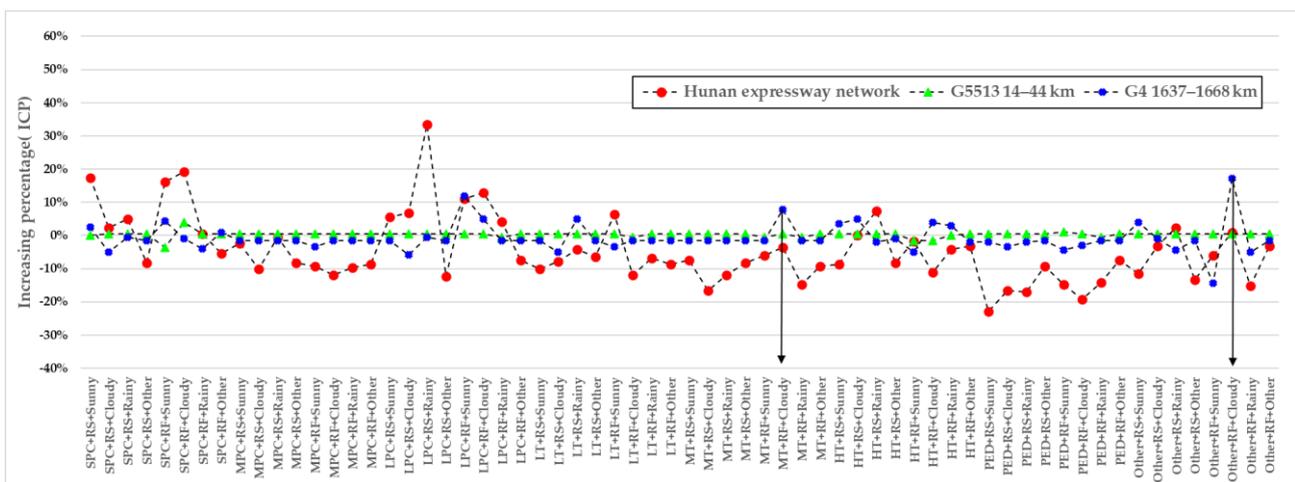
As shown in Figure 14, we make a visual analysis of ICP values under different conditional combinations for the whole expressway and two black spots. In general, the whole expressway network and two black spots show diverse ICP distributions of various conditional combinations, of which the whole expressway network is more fluctuant, but the two black spots are relatively stable. Under most conditional combinations, black spot G5513 14-44 km always keeps low ICP values close to 0%, while black spot G4 1637-1668 km maintains a relatively high ICP value for simple accidents (5%), minor accidents (-10%), and extra-serious (10%) accidents. Moreover, we plot some black arrowed lines to point to some factor combinations with high ICP values for the two black spots, which should be paid more attention to for safety improvement.



(a) 1st Simple Accident

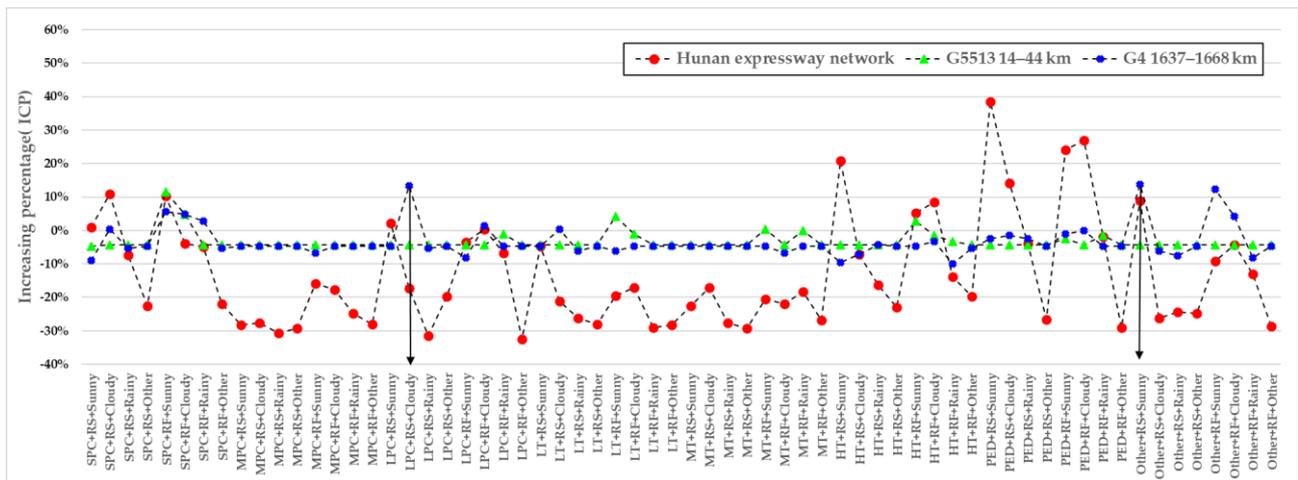


(b) 2nd Minor Accident

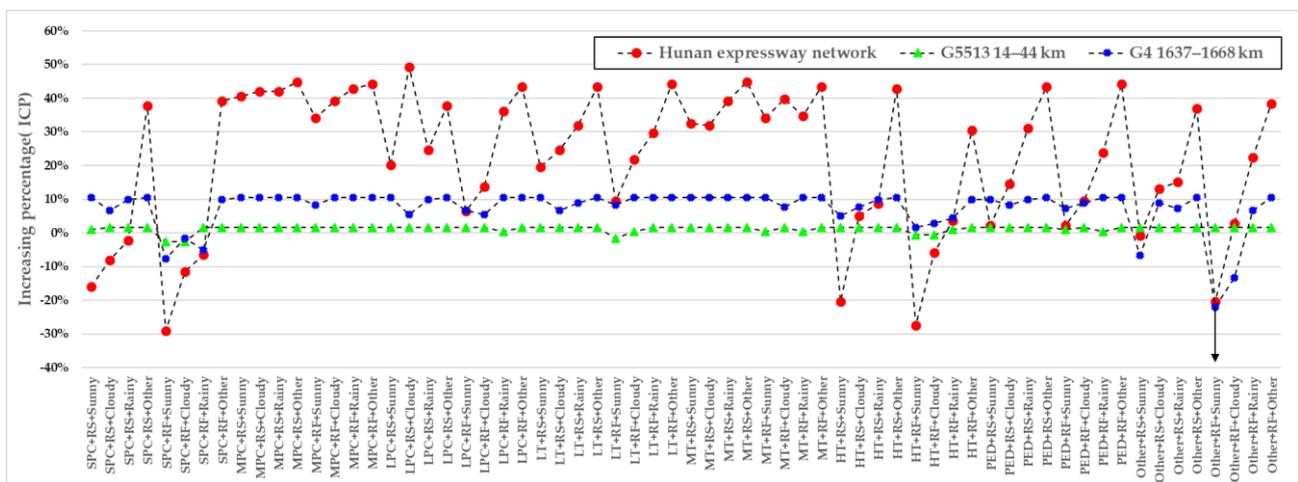


(c) 3rd Ordinary Accident

Figure 14. Cont.



(d) 4th Major Accident



(e) 5th Extra-Serious Accident

**Figure 14.** Increasing percentage (ICP) of conditional probabilities under different combination of vehicle types, weather, and road types (SPC/MPC/LPC: Small/Middle/Large Passenger Car; LT/MT/HT: Light/Middle/Heavy Truck; PED: Pedestrian; other: other traffic modes; RS: Road Sections; RF: Road Facilities).

### 5. Conclusions

Expressways often have high driving speeds, diverse vehicle types, and isolated traffic environments, which make it more likely for serious traffic accidents to happen and cause massive traffic jams. Exploring the spatiotemporal patterns of traffic accident data is one essential prerequisite for discovering expressway safety deficiencies and then making valid countermeasures. Previous studies made significant contributions to crash data analysis but still have some problems that need to be solved, including optimal segmentation of basic statistical units, spatiotemporal aggregation of accident-prone black spots, and spatiotemporal heterogeneities of accident causal factors. Hence, we first introduced temporal periodicity and accident severity into the DBSCAN clustering algorithm for adaptively discovering multi-scale expressway black spots. Then, we applied a sophisticated Bayesian Network model [36] and a divide-and-conquer inference strategy to explore accident causal factors for both local-scale black spots and regional-scale expressway networks. 2012–2016 expressway accident data in Hunan province, China is used to identify accident-prone black spots and explore heterogeneous accident causal factors, showing that the proposed method can correctly extract spatiotemporal aggregation patterns of multi-scale expressway

black spots and meanwhile efficiently discover diverse causal factors for various black spots and expressways. However, there are still some imperfect problems to be negotiated in future research. The first is collaboratively modeling both temporal periodicity and time sequences for analyzing the spatiotemporal evolution of accident-prone black spots. The second is integrating traffic flow data and social-economic data into accident causal analysis for understanding the mechanism of traffic accidents comprehensively.

**Author Contributions:** Conceptualization, Yunfei Zhang and Qiuping Li; Methodology, Yunfei Zhang and Fangqi Zhu; Validation, Fangqi Zhu and Yajun Xie; Investigation, Zehang Qiu; Writing—original draft, Yunfei Zhang and Yajun Xie; Writing—review and editing, Qiuping Li; Visualization, Zehang Qiu. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Nature Science Foundation of China [grant number 41971421 and 41601495], The science and technology innovation Program of Hunan [grant number 2021RC3099] and Natural Science Foundation of Hunan Province [grant number 2022JJ30590].

**Data Availability Statement:** Data is unavailable due to privacy and ethical restrictions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Hazaymeh, K.; Almagbile, A.; Alomari, A.H. Spatiotemporal Analysis of Traffic Accidents Hotspots Based on Geospatial Techniques. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 260. [CrossRef]
- Road Traffic Injuries. Available online: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries> (accessed on 12 December 2022).
- China Statistical Yearbook. Available online: <http://www.stats.gov.cn/tjsj/ndsj/2021/indexeh.htm> (accessed on 12 December 2022).
- Zheng, Z.; Wang, Z.; Zhu, L.; Jiang, H. Determinants of the congestion caused by a traffic accident in urban road networks. *Accid. Anal. Prev.* **2020**, *136*, 105327. [CrossRef]
- Marshall, W.E.; Garrick, N.W. Does street network design affect traffic safety? *Accid. Anal. Prev.* **2011**, *43*, 769–781. [CrossRef]
- Ulak, M.B.; Ozguven, E.E.; Vanli, O.A.; Horner, M.W. Exploring alternative spatial weights to detect crash hotspots. *Comput. Environ. Urban Syst.* **2019**, *78*, 101398. [CrossRef]
- Li, Y.; Bai, Y. Comparison of characteristics between fatal and injury accidents in the highway construction zones. *Saf. Sci.* **2008**, *46*, 646–660. [CrossRef]
- Traffic Management Bureau of the Ministry of Public Security. *Statistical Annual Report of Road Traffic Accidents of the People's Republic of China in 2016*; Traffic Management Bureau of the Ministry of Public Security: Beijing, China, 2017.
- Fang, S.E.; Guo, Z.Y.; Yang, Z. A new identification method for accident prone location on highway. *J. Traffic Transp. Eng.* **2001**, *1*, 90.
- Ouni, F.; Belloumi, M. Pattern of road traffic crash hot zones versus probable hot zones in Tunisia: A geospatial analysis. *Accid. Anal. Prev.* **2019**, *128*, 185. [CrossRef]
- Jorgensen, R.; Research, I. *Evaluation of Criteria for Safety Improvements on the Highway*; Report to the Bureau of Public Roads: Washington, DC, USA, 1966.
- Hauer, E. *Observational Before—After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*; Pergamon Press (Elsevier Science): Oxford, UK, 1997.
- Deacon, J.A.; Zeeger, C.V.; Deen, R.C. Identification of hazardous rural highway locations. *Transp. Res. Rec.* **1975**, *543*, 16–33.
- Thomas, I. Spatial data aggregation: Exploratory analysis of road accidents. *Accid. Anal. Prev.* **1996**, *28*, 251–264. [CrossRef] [PubMed]
- Steenberghen, T.; Dufays, T.; Thomas, I.; Flahaut, B. Intra-urban location and clustering of road accidents using GIS: A Belgian example. *Int. J. Geogr. Inf. Sci.* **2004**, *18*, 169–181. [CrossRef]
- Lipovac, K.; Jovanovi, D.; Bai, S. Mapping of risks on the main road network of Serbia. *Ann. Fac. Eng. Hunedoara* **2009**, *7*, 125.
- Bíl, M.; Andrášik, R.; Sedoník, J. Which curves are dangerous? A network-wide analysis of traffic crash and infrastructure data. *Transp. Res. Part A Policy Pract.* **2019**, *120*, 252–260. [CrossRef]
- Krueger, R.; Bansal, P.; Buddhavarapu, P. A new spatial count data model with Bayesian additive regression trees for accident hot spot identification. *Accid. Anal. Prev.* **2020**, *144*, 105623. [CrossRef]
- Debrabant, B.; Halekoh, U.; Bonat, W.H.; Hansen, D.L.; Hjelmberg, J.; Lauritsen, J. Identifying traffic accident black spots with Poisson-Tweedie models. *Accid. Anal. Prev.* **2018**, *111*, 147–154. [CrossRef] [PubMed]
- Li, L.; Zhu, L.; Sui, D.Z. A GIS-based Bayesian approach for analyzing spatial-temporal patterns of intra-city motor vehicle crashes. *J. Transp. Geogr.* **2007**, *15*, 274–285. [CrossRef]
- Ramírez, A.F.; Valencia, C. Spatiotemporal correlation study of traffic accidents with fatalities and injuries in Bogota (Colombia). *Accid. Anal. Prev.* **2021**, *149*, 105848. [CrossRef]

22. Anderson, T.K. Kernel density estimation and K-means clustering to profile road accident hotspots. *Accid. Anal. Prev.* **2009**, *41*, 359–364. [[CrossRef](#)] [[PubMed](#)]
23. Eckley, D.C.; Curtin, K.M. Evaluating the spatiotemporal clustering of traffic incidents. *Comput. Environ. Urban Syst.* **2013**, *37*, 70–81. [[CrossRef](#)]
24. Blazquez, C.A.; Picarte, B.; Calderón, J.F.; Losada, F. Spatial autocorrelation analysis of cargo trucks on highway crashes in Chile. *Accid. Anal. Prev.* **2018**, *120*, 195–210. [[CrossRef](#)]
25. Harirforoush, H.; Bellalite, L. A new integrated GIS-based analysis to detect hotspots: A case study of the city of Sherbrooke. *Accid. Anal. Prev.* **2019**, *130*, 62–74. [[CrossRef](#)]
26. Holmgren, J.; Knapen, L.; Olsson, V.; Masud, A.P. On the use of clustering analysis for identification of unsafe places in an urban traffic network. *Procedia Comput. Sci.* **2020**, *170*, 187–194. [[CrossRef](#)]
27. Xu, P.; Huang, H.; Dong, N. The modifiable areal unit problem in traffic safety: Basic issue, potential solutions and future research. *J. Traffic Transp. Eng. (Engl. Ed.)* **2018**, *5*, 73–82. [[CrossRef](#)]
28. Ghadi, M.; Török, Á. A comparative analysis of black spot identification methods and road accident segmentation methods. *Accid. Anal. Prev.* **2019**, *128*, 1–7. [[CrossRef](#)]
29. Black, W.R. Highway accidents: A spatial and temporal analysis. *Transp. Res. Rec.* **1991**, *1318*, 75–82.
30. Golze, J.; Feuerhake, U.; Koetsier, C.; Sester, M. Impact analysis of accidents on the traffic flow based on massive floating car data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2021**, *XLIII-B4-202*, 95–102. [[CrossRef](#)]
31. Zhang, G.; Yau, K.K.W.; Zhang, X.; Li, Y. Traffic accidents involving fatigue driving and their extent of casualties. *Accid. Anal. Prev.* **2016**, *87*, 34–42. [[CrossRef](#)]
32. Thiffault, P.; Bergeron, J. Monotony of road environment and driver fatigue: A simulator study. *Accid. Anal. Prev.* **2003**, *35*, 381–391. [[CrossRef](#)]
33. Pervez, A.; Huang, H.; Lee, J.; Han, C.; Wang, J.; Zhang, X. Crash analysis of expressway long tunnels using a seven-zone analytic approach. *J. Transp. Saf. Secur.* **2021**, *13*, 108–122. [[CrossRef](#)]
34. Qu, X.; Zhu, X.; Xiao, X.; Wu, H.; Guo, B.; Li, D. Exploring the Influences of Point-of-Interest on Traffic Crashes during Weekdays and Weekends via Multi-Scale Geographically Weighted Regression. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 791. [[CrossRef](#)]
35. Hong, J.; Tamakloe, R.; Park, D. Application of association rules mining algorithm for hazardous materials transportation crashes on expressway. *Accid. Anal. Prev.* **2020**, *142*, 105497. [[CrossRef](#)]
36. Karimnezhad, A.; Moradi, F. Road accident data analysis using Bayesian networks. *Transp. Lett.* **2017**, *9*, 12–19. [[CrossRef](#)]
37. Hwang, S.; Boyle, L.N.; Banerjee, A.G. Identifying characteristics that impact motor carrier safety using Bayesian networks. *Accid. Anal. Prev.* **2019**, *128*, 40–45. [[CrossRef](#)] [[PubMed](#)]
38. 2021 Hunan Statistical Yearbook. Available online: <http://222.240.193.190/2021tjnj/indexeh.htm> (accessed on 12 December 2022).
39. Bíl, M.; Andrášik, R.; Sedoník, J. A detailed spatiotemporal analysis of traffic crash hotspots. *Appl. Geogr.* **2019**, *107*, 82–90. [[CrossRef](#)]
40. Deng, M.; Liu, Q.; Cheng, T.; Shi, Y. An adaptive spatial clustering algorithm based on Delaunay triangulation. *Comput. Environ. Urban Syst.* **2011**, *35*, 320–332. [[CrossRef](#)]
41. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with Noise. In Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, 2–4 August 1996; pp. 226–231.
42. Geng, C.; Peng, Y. Identification method of traffic accident black spots based on dynamic segmentation and DBSCAN algorithm. *J. Chang. Univ. Nat. Sci. Ed.* **2018**, *38*, 131–138.
43. Shinde, P.; Kudalkar, G.; Pawar, S.; Tiwari, H.; Khairnar, H. Khairnar Accident Hotspot Detection by DB-Scan Clustering. In *ICT Analysis and Applications; Lecture Notes in Networks and Systems*; Fong, S., Dey, N., Joshi, A., Eds.; Springer: Singapore; Volume 314.
44. Notice of the Ministry of Public Security on Revising the Classification Standards for Road Traffic Accidents. Available online: <http://fgcx.bjcourt.gov.cn:4601/law?fn=chl026s135.txt> (accessed on 10 January 2023).
45. Antão, P.; Soares, C.G. Analysis of the influence of human errors on the occurrence of coastal ship accidents in different wave conditions using Bayesian Belief Networks. *Accid. Anal. Prev.* **2019**, *133*, 105262. [[CrossRef](#)]
46. Friedman, N.; Dan, G.; Goldszmidt, M. Bayesian Network Classifiers. *Mach. Learn.* **1997**, *29*, 131–163. [[CrossRef](#)]
47. Ni, P.; Yu, W. A Victim-Based Framework for Telecom Fraud Analysis: A Bayesian Network Model. *Comput. Intell. Neurosci.* **2022**, *2022*, 7937355. [[CrossRef](#)]
48. The Ministry of Public Security Announced the Top 10 Black Spots in 2017. Available online: [http://www.gov.cn/xinwen/2018-02/02/content\\_5263057.htm](http://www.gov.cn/xinwen/2018-02/02/content_5263057.htm) (accessed on 12 December 2022).
49. 33 Accident-Prone Sections for Hunan Expressway, See Where Are All? Available online: [https://www.sohu.com/a/195396550\\_99977726](https://www.sohu.com/a/195396550_99977726) (accessed on 16 January 2023).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.