sustainability

MDPI

Article

Identification Method for Crash-Prone Sections of Mountain Highway under Complex Weather Conditions

Rishuang Sun [1,2], Chi Zhang [1,3,*], Yujie Xiang [1,3], Lei Hou [4] and Bo Wang [1,3]

1 School of Highway, Chang'an University, Xi'an 710064, China
2 Shandong Provincial Communication Planning and Design Institutions Group Co., Ltd., Jinan 250031, China
3 Engineering Research Center of Highway Infrastructure Digitalization, Ministry of Education, Xi'an 710064, China
4 School of Engineering, RMIT University, Melbourne, VIC 3000, Australia
* Correspondence: zhangchi@chd.edu.cn

Abstract: Mountain highway crashes usually have a weather tendency, and the crash-prone sections show obvious weather differences. However, there were few targeted quantitative analyses of the impact of weather conditions on crash-prone sections in previous studies. Aiming at the problem that traditional identification methods ignore the difference in weather, this paper proposed the time-spatial density ratio method. The method quantified the length of the road section, the period, and the influence of different weather conditions through the time-spatial density ratio. Then the time-spatial density ratios under different weather conditions were comprehensively sorted in parallel. Finally, the risk threshold was determined according to the characteristics of the cumulative frequency curve's double inflection points, and the crash-prone sections under each weather condition were identified. This paper evaluated the crash-prone sections of the G76 Expressway. Moreover, the crash risk situation under each weather condition was characterized through kernel density analysis. The method was compared with the cumulative frequency method, a traditional method suitable for Chinese highways with similar application conditions. The effective search index was utilized as a comparison factor. The results showed that the effective search index of the time-spatial density ratio method was more than 80% greater than that of the cumulative frequency method.

Keywords: traffic safety; crash-prone sections; time-spatial density ratio method; mountain highway; complex weather

check for updates

Citation: Sun, R.; Zhang, C.; Xiang, Y.; Hou, L.; Wang, B. Identification Method for Crash-Prone Sections of Mountain Highway under Complex Weather Conditions. *Sustainability* **2022**, *14*, 15181. https://doi.org/10.3390/su142215181

Academic Editor: Lei Zhang

Received: 15 October 2022
Accepted: 11 November 2022
Published: 16 November 2022

1. Introduction and Review

Traffic safety is a globally acknowledged issue; traffic crashes are the world's eighth largest cause of death. Annual Report on Road Traffic Accident Statistics shows 265,204 road traffic crashes in China in 2019 [1]. The Decade of Action for Road Safety campaign launched by the World Health Organization set a goal that by 2030, global road traffic fatalities should be reduced by 50% [2]. Due to complex climatic factors, mountain highways often have potential safety hazards in some sections. The number and characteristics of traffic crashes in some locations are significantly more prominent than in other locations, defined as crash-prone sections [3]. Determining the location of crash-prone sections on the road is crucial to enhancing the traffic safety of the road.

There is extensive literature focusing on the identification of crash-prone sections. Among the common identification methods of crash-prone sections, Florida's traffic management department first applied the crash number method in identifying crash-prone points. Since then, the crash number method, the crash frequency method, and the equivalent crash frequency method have been extensively used [4–6]. These methods were typically used as basic methods and combined with other methods to improve the efficiency and accuracy of identification. Jordan [7] proposed the matrix method that comprehensively considered the number of crashes and the crash rate. Fang et al. [3] applied the statistical principle to propose the cumulative frequency curve method to identify crash-prone sections. In addition, the research methods developed include the safety factor method, the equivalent crash number method, etc. [8]. These methods carry out simple statistical analysis based on historical data, and determine the threshold value subjectively or objectively as the identification criterion for crash-prone sections. The traditional identification methods of crash-prone sections have been developed and matured, which are simple, intuitive, and easy to calculate. Nevertheless the analysis of crash characteristics is insufficient because only a few factors were considered.

With the development of related research, clustering analysis methods such as the K-means algorithm, density-based clustering algorithm, grey correlation analysis, and fuzzy clustering have been widely used. Shen [9] established the principal component clustering analysis model according to the cumulative equation contribution rate and component index value. Wan Y et al. [10] proposed a method that can quickly identify and classify crash black spots on urban roads, to solve the problems caused by the randomness of crash occurrences and the unclear classification of crash black spots by the traditional model. Yakar [11] proposed the crash-prone road section identification method based on multicriteria decision-making, and the identification rate can reach more than 67%. Zhang et al. [12] proposed an improved K-means clustering algorithm to solve the shortcomings of the traditional algorithm. The clustering analysis methods can make up for the omission of identification caused by the simple division of road sections, and more objectively reflect the actual length of crash-prone sections. However, the determination of cluster numbers is more subjective, directly affecting the threshold's determination. These methods have certain applicability and limitations.

In order to study the characteristics of traffic crashes, many statistical methods were used in crash analysis, such as Poisson distribution, negative binomial distribution, Bayesian model, etc. To study the relationship between the road and crash-prone sections, Wang J et al. [13] quantitatively examined the effect of 3D road alignment on traffic safety on mountainous freeways by the Bayesian Tobit model. Since the traffic volume and the crash rate are not directly proportional, Saccomanno et al. [14] indicated that the Bayesian model could be more accurately applied to road sections with complex traffic compositions. Malyshkina and Mannering [15] proposed a Markov transformation counting data model to predict the number of crashes on the road. Barua et al. [16] applied a stochastic effect model considering heterogeneity to study spatial autocorrelation in traffic crashes. Debrabant et al. [17] used statistical methods such as Poisson-Tweedie distribution to identify the black spots of crashes. Methods based on mathematical models have the characteristics of higher identification accuracy. However, most methods are based on specific model assumptions and predefined potential relationships between dependent and independent variables. If these assumptions are violated, some models may mistakenly identify crash-prone sections [18].

Based on the shortcomings of traditional methods, new methods combined with machine learning algorithms and spatial analysis techniques have been vigorously developed in recent years. Zhang et al. [19] investigated stratification heterogeneity in Shenzhen traffic crashes, what factors influence the casualties, and the interaction of those factors. Fan et al. [20] proposed a black-spot recognition algorithm based on the support vector machine and a black-spot analysis method based on the deep neural network. Gutierrez-Osorio and Pedraza [21] reviewed algorithms and models including data mining and machine learning techniques for analyzing, characterizing, and predicting road crashes. These methods find a new way to identify crash-prone sections and improve the shortcomings of traditional methods. Nevertheless, each method has its applicability and the related research on new methods has yet to establish a system.

Given the above, most studies on the identification of crash-prone sections are based on the spatial discreteness of traffic crashes. The depth of the potential causes of crashes is insufficient, especially the relationship between complex weather conditions and crashes. The weather predominantly affects many traffic crashes, especially in places with complex weather. The impact of weather on crashes is significant. Many methods average out the impact of weather on crashes. In this area of research, Ahmed et al. [22] used the Bayesian hierarchy method to study the impact of special weather on mountain highways. Yu et al. [23] used the Bayesian inference method to obtain weather condition variables, especially precipitation, which played a crucial role in the crash model. Yakar [24] considered the spatial attributes of crashes and established the relationship between the number of crashes and road environmental characteristics. The abovementioned studies considered the crash distribution from the road environment, weather conditions, meteorological indicators, and other factors, but did not quantify their impact on the crash distribution.

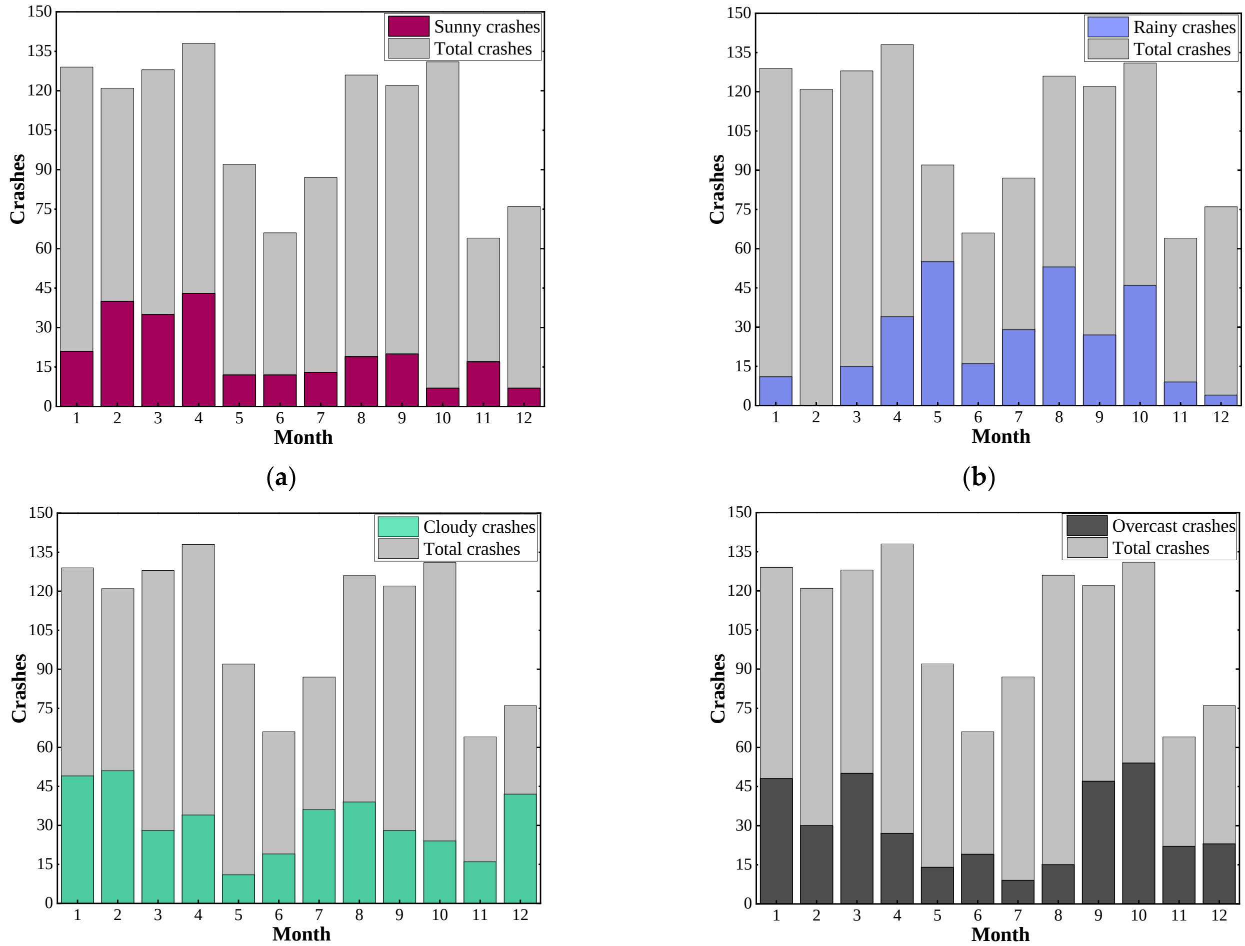
Based on the overall literature review, few studies have investigated the quantitative methodologies of the impact of different weather conditions on crash-prone sections. This research aimed to study the identification of crash-prone sections under complex weather conditions by establishing the relationship between the number of crashes and the time-spatial characteristics of crashes.

2. Data Preparation

The object of this study is the G76 Expressway in the southwest mountainous area of China, with a total length of 135.875 km, as shown in Figure 1. The subgrade of the expressway is 24.5 m wide, and it is a two-way four-lane road with a design speed of 80 km/h. The high-pier bridges and long tunnels are densely constructed, with many long longitudinal slope sections. It has distinctive features of complex topography, geology, and relatively harsh weather conditions.

Figure 1. Study area.

Determining the analysis period is very important. From a purely statistical point of view, a high number of crashes benefits identification accuracy. On the other hand, within the analysis period, many changes in the field (traffic flow, road conditions, traffic policies, vehicle proportions, etc.) will affect the analysis results, limiting the selection of the period. As a result, the period for identifying crash-prone sections ranges from 1 to 5 years (Yakar [24]). In this case, the traffic control situation is complex. The crash data from January 2018 to April 2021 were used for analysis to ensure that the driving environment would not change much during the period. During the study period, there were 1286 traffic crash data records, and the crash situation under each weather condition is shown in Figure 2. In addition to the picture, there were two foggy crashes in January and three in February.



(**a**) (**b**) (**c**) (**d**)

Figure 2. Statistics of crash data under (**a**) sunny, (**b**) rainy, (**c**) cloudy, and (**d**) overcast weather.

A crucial step in data preparation is road segmentation. The determination of the unit length is critical in the identification of crash-prone sections. The fixed-length method makes it difficult to indicate what the optimal length of the segment should be. If the selected length is too long, it is difficult to ensure the authenticity and accuracy of the unit crash situation. On the other hand, if the length is too short, it may result in insufficient data precision (Yakar [24]). These issues should be taken into account when dividing roads. Maen et al. [25] compared the effect of methodological diversity of road network segmentation on the performance of different crash-prone sections identification methods. A study documented the performance evaluation results of three different highway segmentation methods (Kwon et al. [26]). In this study, the homogeneity method was used to take the change point of the highway's horizontal and vertical linear elements as the dividing point. The road was divided into linear fixed road units, and each unit was numbered as shown in Figure 3. The research road was divided into 677 units, and the number of crashes was matched for each unit. A database was established according to the division unit combined with the historical weather data of the crash, as shown in Figure 4.

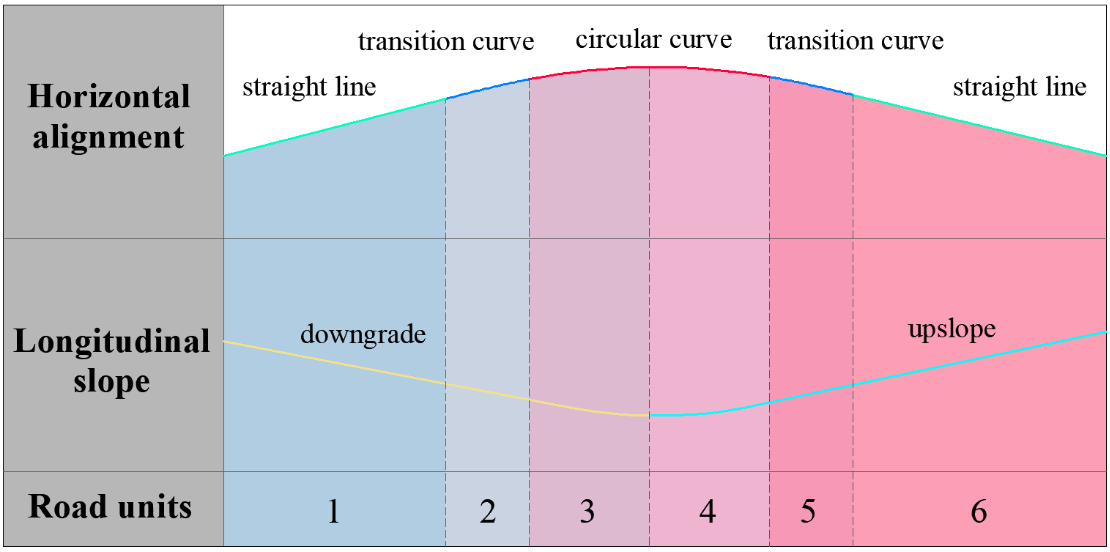
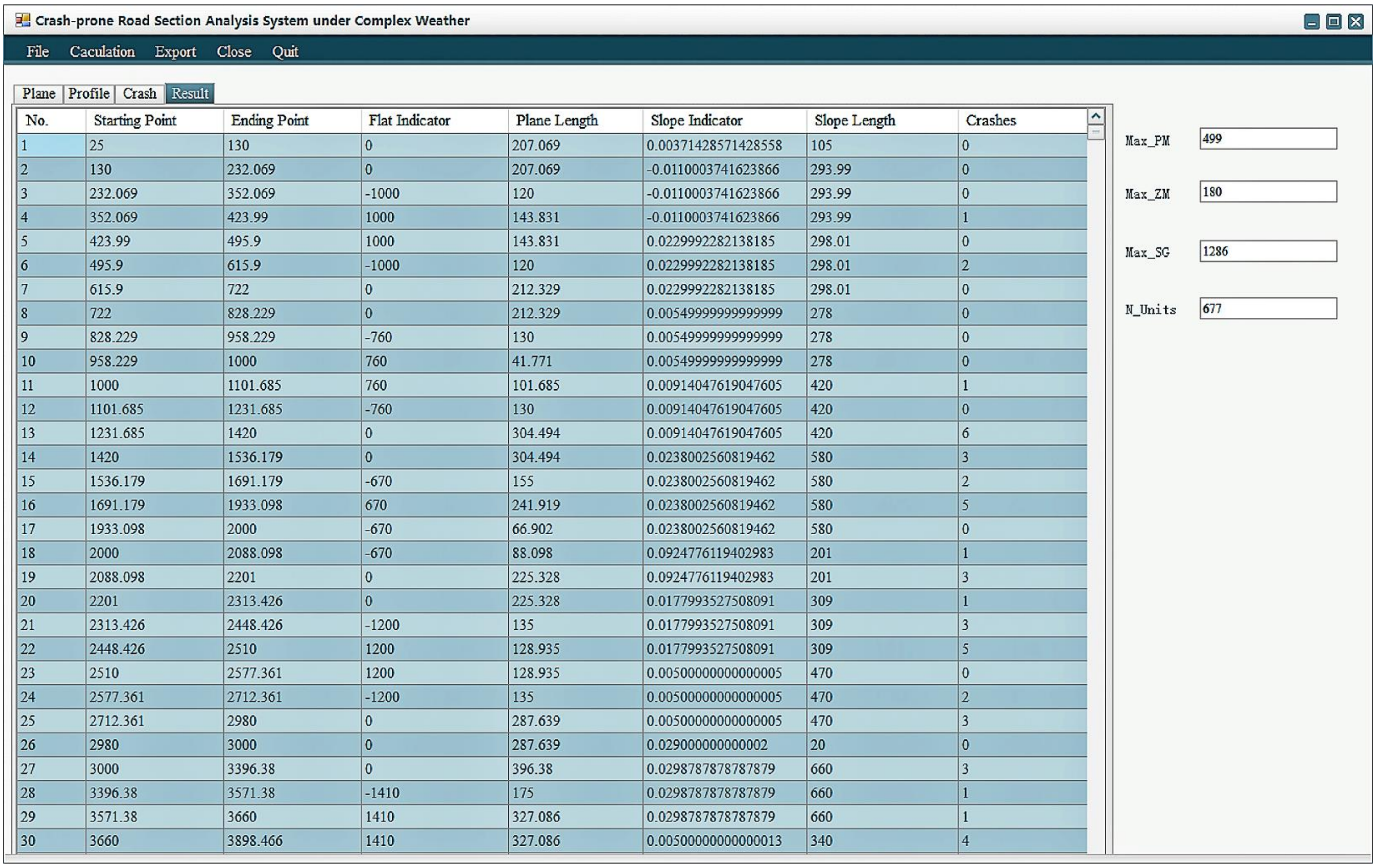


Figure 3. Section division by homogeneous method.



No.	Starting Point	Ending Point	Flat Indicator	Plane Length	Slope Indicator	Slope Length	Crashes
1	25	130	0	207.069	0.00371428571428558	105	0
2	130	232.069	0	207.069	-0.0110003741623866	293.99	0
3	232.069	352.069	-1000	120	-0.0110003741623866	293.99	0
4	352.069	423.99	1000	143.831	-0.0110003741623866	293.99	1
5	423.99	495.9	1000	143.831	0.0229992282138185	298.01	0
6	495.9	615.9	-1000	120	0.0229992282138185	298.01	2
7	615.9	722	0	212.329	0.0229992282138185	298.01	0
8	722	828.229	0	212.329	0.00549999999999999	278	0
9	828.229	958.229	-760	130	0.00549999999999999	278	0
10	958.229	1000	760	41.771	0.00549999999999999	278	0
11	1000	1101.685	760	101.685	0.00914047619047605	420	1
12	1101.685	1231.685	-760	130	0.00914047619047605	420	0
13	1231.685	1420	0	304.494	0.00914047619047605	420	6
14	1420	1536.179	0	304.494	0.0238002560819462	580	3
15	1536.179	1691.179	-670	155	0.0238002560819462	580	2
16	1691.179	1933.098	670	241.919	0.0238002560819462	580	5
17	1933.098	2000	-670	66.902	0.0238002560819462	580	0
18	2000	2088.098	-670	88.098	0.0924776119402983	201	1
19	2088.098	2201	0	225.328	0.0924776119402983	201	3
20	2201	2313.426	0	225.328	0.0177993527508091	309	1
21	2313.426	2448.426	-1200	135	0.0177993527508091	309	3
22	2448.426	2510	1200	128.935	0.0177993527508091	309	5
23	2510	2577.361	1200	128.935	0.00500000000000005	470	0
24	2577.361	2712.361	-1200	135	0.00500000000000005	470	2
25	2712.361	2980	0	287.639	0.00500000000000005	470	3
26	2980	3000	0	287.639	0.029000000000002	20	0
27	3000	3396.38	0	396.38	0.0298787878787879	660	3
28	3396.38	3571.38	-1410	175	0.0298787878787879	660	1
29	3571.38	3660	1410	327.086	0.0298787878787879	660	1
30	3660	3898.466	1410	327.086	0.00500000000000013	340	4

Figure 4. Database creation.

3. Methodology

3.1. Framework for Research Design

The methodology of applying crash data analysis is shown in in Figure 5.

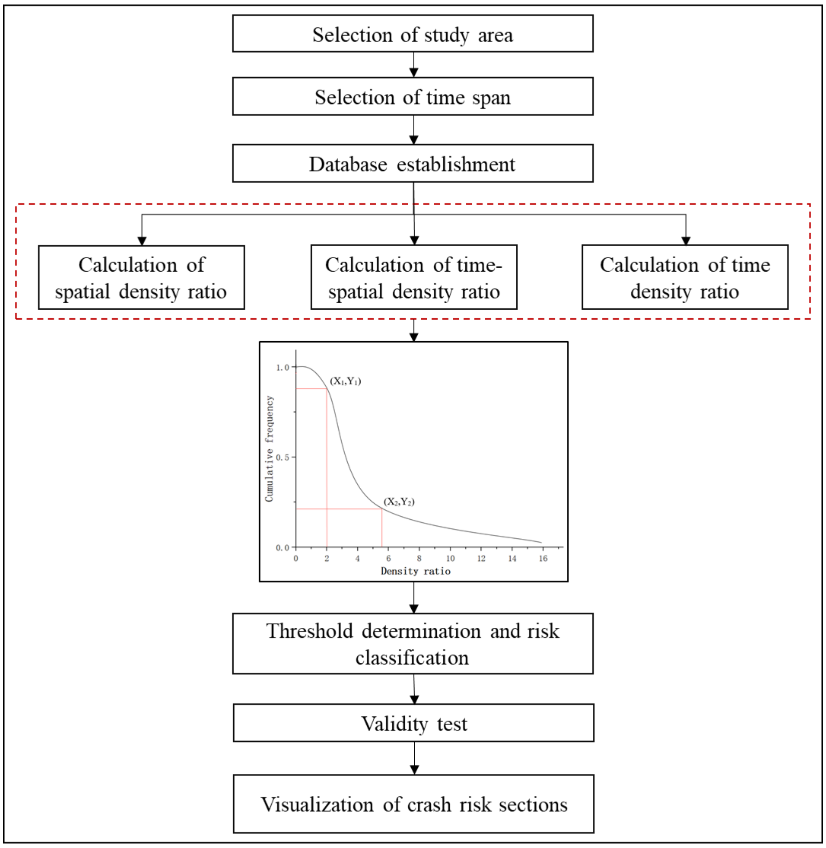


Figure 5. Methodological framework.

Step 1. Determination of study area and time span.
Step 2. Linear unit division and database establishment.
Step 3. Calculation of the density ratio.
Step 4. Application of cumulative frequency method to determine the risk threshold and grading by the characteristic of the cumulative frequency curve.
Step 5. Validity testing of identified crash-prone sections.
Step 6. Visual understanding of highway crash risk areas using AcrGIS's kernel density analysis.

3.2. Spatial and Time Density Ratio Indicators

Spatial density expresses the crash rate measured by length. The purpose is to balance the effect of the length of the road segment unit on the crash distribution. The number of days for different weathers in the study period varied widely. Time density expresses the crash rate measured by time. The purpose is to eliminate the error of the number of days in different weather to study the risk of different weather.

The basic theme of spatial and time density ratio indicators is to characterize the relative safety level of road sections. Yakar [24] proposed similar relative analytical indicators but lacked meteorological considerations. These indicators can be used in various ways according to the data source situation and analysis requirements representing the crash risk of different dimensions. The weather-specific crash risk identification indicators were proposed by distinguishing crash distribution under different weather conditions.

3.2.1. Spatial Density Ratio

From a spatial perspective, the influence of the length of units on the crash frequency is considered, which is quantified by calculating the crash spatial density. According to the ratio of the unit crash spatial density to the road crash spatial density, the spatial density ratio is obtained as shown in Equation (1), representing the relative value of the crash spatial distribution.

$$P_s = \frac{N_e/L_e}{N/L} \tag{1}$$

where P_s is the spatial density ratio, L_e is the unit length, as shown in the fifth column of data in Figure 3. N_e is the number of crashes in the unit, as shown in the eighth column of data in Figure 3. N is the number of crashes on the road, $N = \sum N_e$. L is the length of the road, $L = \sum L_e$.

3.2.2. Time–Density Ratio

The influence of complex weather conditions and period on crash frequency is considered from a time perspective. For each weather condition, crash time densities were calculated to quantify the effect of the period. The unit time density ratio and the road time density ratio are obtained as shown in Equations (2) and (3). The two indicators respectively represent the relative value of the accident time distribution of each unit and the road under each weather condition.

$$P_{t_e} = \frac{N_{x_e}/D_x}{N_e/D} \tag{2}$$

$$P_t = \frac{N_x/D_x}{N/D} \tag{3}$$

where P_{t_e} is the unit time density ratio, P_t is the road time density ratio, N_{x_e} is the number of crashes in the unit in x weather, N_x is the number of crashes on the road in x weather, $N_x = \sum N_{x_e}$. D_x is the number of days in x weather, D is the total number of days, $D = \sum D_x$.

3.2.3. Time-Spatial Density Ratio

The influence of unit length and time span under complex weather is comprehensively considered from a time-spatial perspective. Its combined effect is quantified by calculating the time-spatial density ratio as shown in Equation (4). The time-spatial density ratio attribute under each weather condition is assigned to each unit to represent the relative time-spatial value of crash risk.

$$P_{ts} = \frac{N_{x_e}/L_e/D_x}{N/L/D} \tag{4}$$

where P_{ts} is the time-spatial density ratio.

3.3. Risk Threshold and Classification

There are relatively few safe sections and crash-prone sections on actual roads, but more low-risk sections. In this study, the cumulative frequency diagram was drawn through the density ratio of each unit. The cumulative frequency curve has the characteristic of double inflection points, as shown in Figure 6. The density ratios corresponding to the double inflection points were used as the crash risk threshold. The risk levels were divided into three categories: level I represented the relatively safe section, level II represented the low-crash section, and level III represented the crash-prone section.

For determining the inflection point, Wei and Wen [27] adopted a relatively objective method of selecting the inflection point of the cumulative frequency curve. All points in the curve were used for linear regression analysis. The curve regression was set to satisfy two conditions: the determinant coefficient (R^2) greater than 0.9 and the significance level (p) less than 0.05. When the conditions were not met, the uppermost point of the curve was gradually discarded. The point that satisfied the conditions is the inflection point. Fan et al. [28] proposed a double inflection points identification method based on this method. Starting from the bottom end of the curve, the determinant coefficient of the fitted curve between each point and the bottom end was calculated sequentially. In the beginning, the determinant coefficient would have a series of fluctuations, and then it would gradually increase to the highest point, the inflection point.

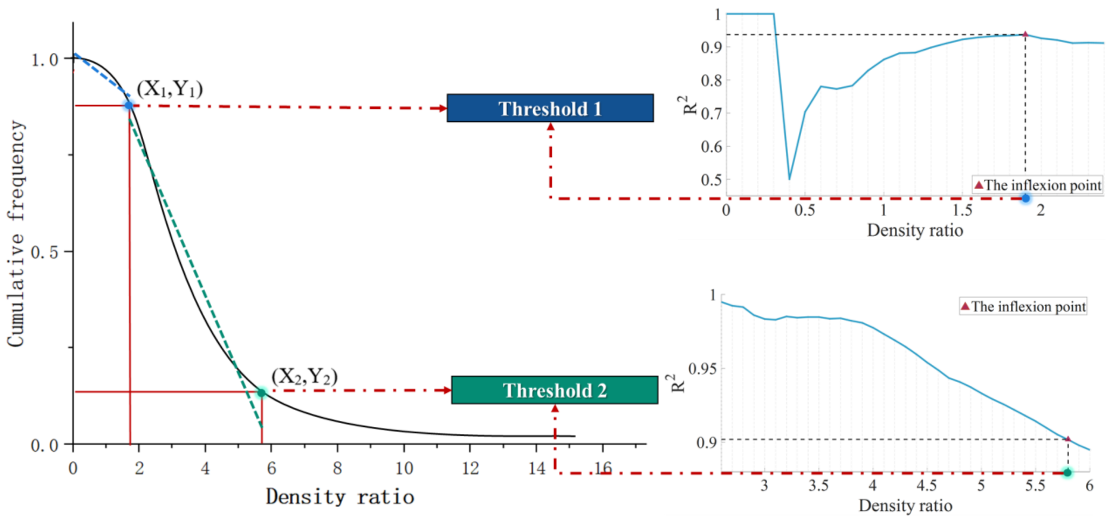


Figure 6. Determination of risk thresholds.

The two abovementioned methods have different applicable curves for selecting inflection points. This study was based on these two methods. According to the characteristic of the cumulative frequency curve of the density ratio, the first inflection point was the position of the maximum determinant coefficient of the fitted straight line. Starting from this inflection point, the determinant coefficient of the fitted straight line between the points after the first point and the first inflexion point was calculated in turn. The point until the determinant coefficient of critical 0.9 was taken as the second inflection point, as shown in Figure 6.

3.4. Kernel Density Analysis

Kernel density analysis is a nonparametric spatial analysis method that calculates the density of a feature in its surrounding neighborhood [29]. According to the number of crashes, Xie et al. [30] integrated NetKDE with native Moran's I for hotspot detection of traffic crashes and evaluated statistical significance using two Monte Carlo simulations. Anderson [31] presented a methodology using Geographical Information Systems and Kernel Density Estimation to study the spatial patterns of injury-related road crashes. Mohaymany et al. [32] used the results of network kernel density estimation for fatal, injury, and property damage-only crashes to estimate the spatial risk pattern of crashes. Due to the complexity of crash causes, crash risk areas are often not only located in crash-prone sections. The crash risk areas of the road can be reflected according to the density ratio through the kernel density analysis method.

In this study, each unit was assigned a density ratio attribute value in Arcmap. The Gaussian kernel function was employed for kernel density analysis based on the traffic safety risk characteristics (Silverman [33]). The result of the kernel density analysis was used to characterize the crash risk situation. The crash risk distribution of the road was displayed in 3D through ArcScene.

3.5. Method Comparison and Test

The time-spatial density ratio method is based on the density ratio indicator to analyze crashes under complex weather conditions. The density ratio of each unit was sorted, and the crash-prone sections were determined based on the threshold. An essential traditional method similar to this method is the cumulative frequency method.

The cumulative frequency method is based on the number of unit crashes, and the crash threshold is determined according to the cumulative frequency curve to judge the crash-prone sections. This method avoids the problem that the unified standard value cannot meet the actual needs of different projects, and is suitable for the actual situation where the driving safety level of each road in China is quite different (Wu et al. [34]).

In this study, the effective search index was proposed to test the validity of the identification results of crash-prone sections. The identification effect was evaluated by the relative proportion of the number of crashes and the length of the road section, as shown in Equation (5). The practicability of the time-spatial density ratio method was verified by comparing the cumulative frequency method.

$$SEI = \frac{N_i/N}{L_i/L} \tag{5}$$

where SEI is the search efficiency index, N_i is the number of crashes on identified sections, and L_i is the length of identified sections.

4. Results

4.1. Spatial Distribution of Crashes

The attribute of the spatial density ratio of each unit was given. The corresponding values of the double inflection points of the cumulative frequency curve calculated by the threshold determination method were 0.5 and 2.8. As indicated in Figure 7, the road risk levels were separated into three categories. According to the cumulative frequency method, the unit with more than five incidents was the crash-prone section. The effective search index was calculated respectively, and the identification effect is shown in Table 1.

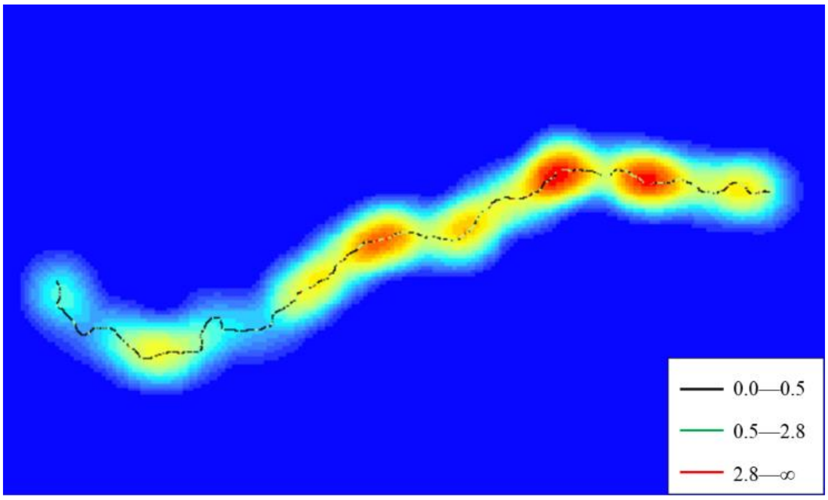


Figure 7. Distribution of spatial density ratio.

Table 1. Identification effect of the spatial density ratio method.

Risk Level	Ps	Length of the Road (m)	Ratio of Identified Length to the Total Length	Number of Crashes	Ratio of Identified Crashes to Total Crashes	Search Efficiency Index	
						Spatial Density Ratio Method	Cumulative Frequency Method
I	0~0.5	58,958.15	0.434	66	0.051	0.12	2.97
II	0.5~2.8	68,263.48	0.502	758	0.589	1.17	
III	2.8~∞	8652.962	0.064	462	0.360	5.64	

Level I sections account for 43.4% of the length, and only 5.1% of crashes occurred. Level III sections account for 6.4% of the length but have 36.0% of crashes. Most crashes occur on a small number of dangerous road sections. The effective search index of the cumulative frequency identification method based on the spatial density ratio is much larger than that based on the number of crashes. The identification of crash-prone sections is more accurate after balancing the length of the road section.

4.2. Time Distribution of Crashes

4.2.1. Time Distribution of Unit Crashes

The time–density ratio attribute of each unit under each weather condition condition was given, which was utilized to distinguish the influence of different weather conditions on the crash risk. The time density ratio distribution of each unit is shown in Figure 8 (the locations marked in the figure are crash-prone sections identified from a spatial perspective). The length of each bar with different colors represented the relative value of crash risk under each weather condition of the unit.

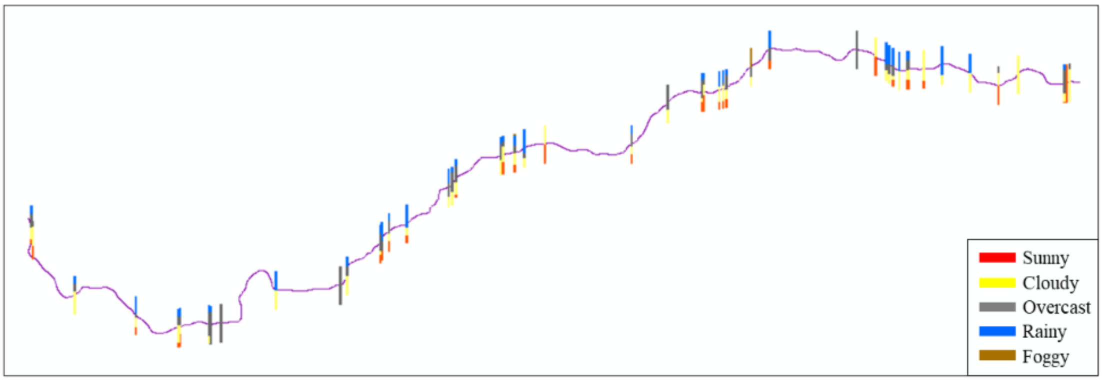


Figure 8. Distribution of unit time–density ratio.

Units that were significantly affected by weather conditions could be sorted out. This can be used as the basis for the targeted security improvement of the unit, as shown in Table 2.

Table 2. Units heavily affected by complex weather conditions.

No.	Number of Crashes	P_{t_e} of Sunny Days	P_{t_e} of Cloudy Days	P_{t_e} of Overcast Days	P_{t_e} of Rainy Days	P_{t_e} of Foggy Days
111	11	0.82	0.91	0.33	2.31	0
308	5	0.9	0.67	0	3.05	0
309	4	0	0.83	2.73	0	0
359	5	2.69	0	0.73	1.02	0
365	4	0	0.83	0	2.54	0
380	8	0	0.83	0.46	3.18	0
384	4	0	0	0.91	3.82	0
385	16	0.84	1.04	0.68	1.27	12.67
			……			
662	4	0	0.83	0	3.82	0

4.2.2. Time Distribution of Road Crashes

The time–density ratio of the road in various weather was given, which was used to characterize the impact of weather conditions on the risk of road crashes. The time–density ratio distribution of the road is shown in Figure 9. It can be seen from the figure that the crash risk of the road in rainy weather was significantly higher than that in other weather. Only six days were foggy within the research period due to the less foggy weather in the study area. Traffic crashes in foggy weather were more occasional, and the management department would take control measures such as closing highways. Therefore, there was no absolute reference value for foggy days. It can reflect the relative safety level of highways under various weather conditions based on sufficient data.

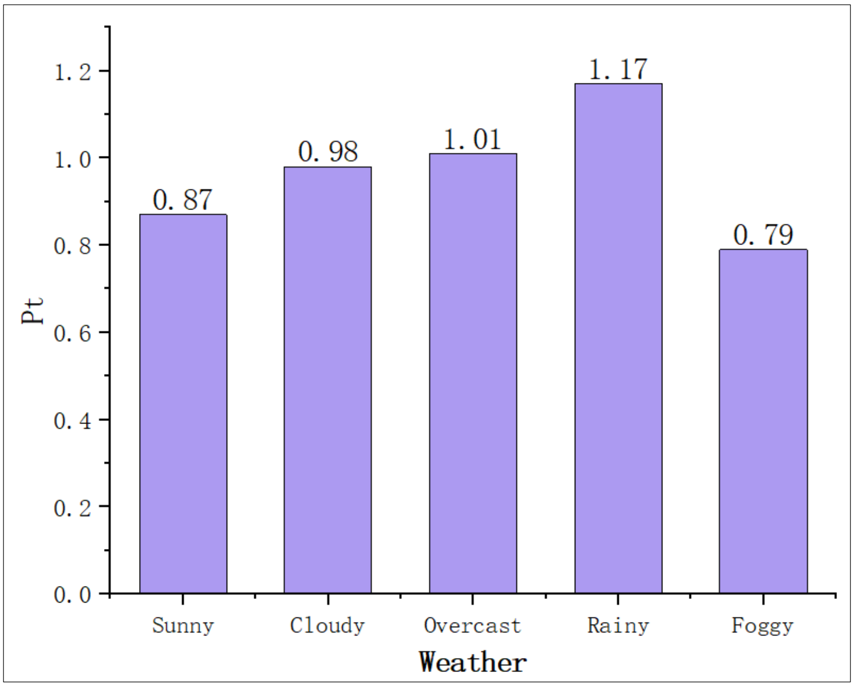


Figure 9. Distribution of road time density ratio.

4.3. Time-Spatial Distribution of Crashes

The time-spatial density ratio of each unit in each weather condition was assigned and sorted in parallel. The corresponding values of double inflection points of the cumulative frequency curve were 1.9 and 5.8, as shown in Figure 10, which were used as the risk threshold to divide the risk level into three categories.

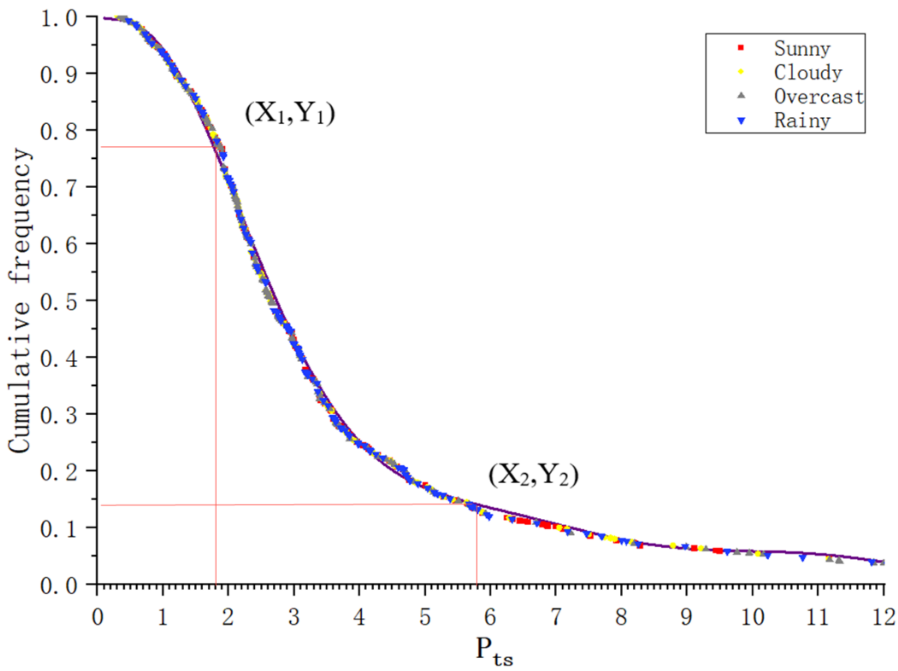


Figure 10. Cumulative frequency curve of time-spatial density ratio.

The crash-prone sections and the crash risk represented by the kernel density under each weather condition are shown in Figure 11.

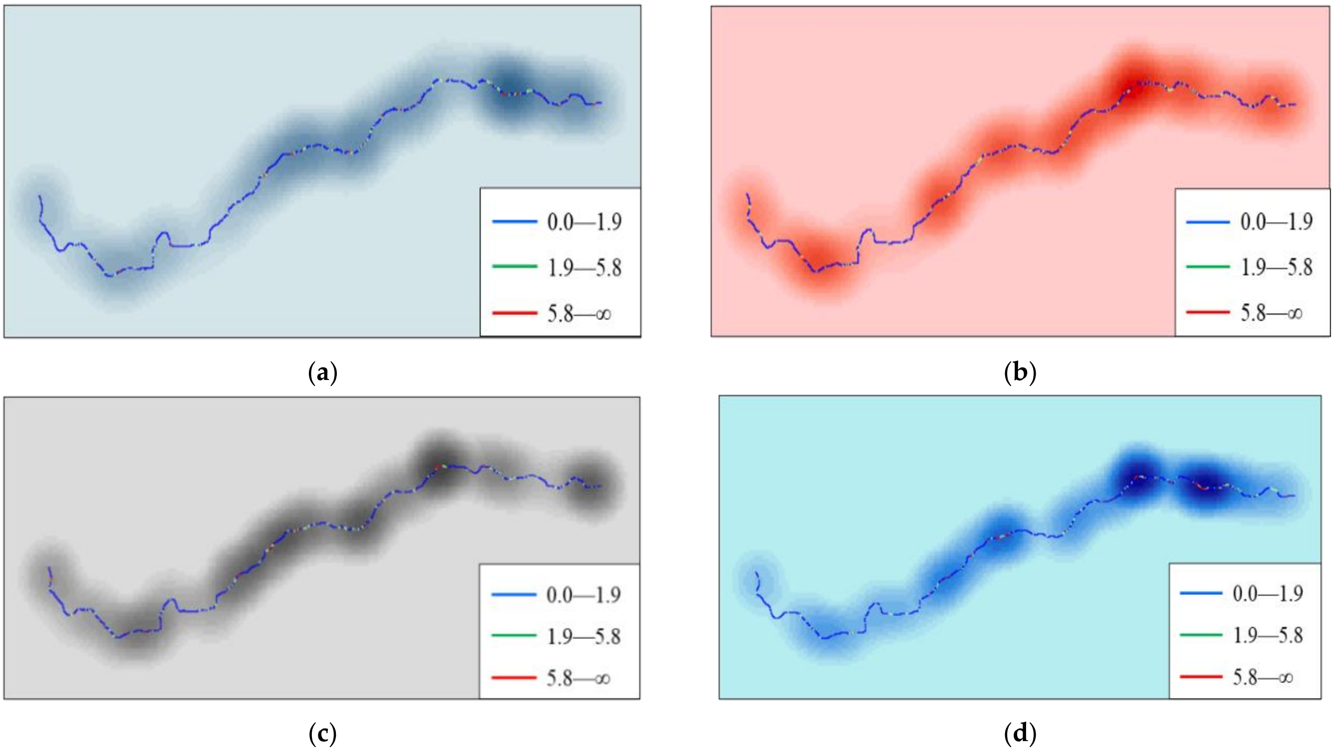


Figure 11. Distribution of time-spatial density ratio in different weather conditions. (**a**) sunny, (**b**) cloudy, (**c**) overcast, and (**d**) rainy.

ArcScene was used to visualize the time-spatial density ratio kernel density values in 3D. The complex-phase road crash risk distribution under complex weather conditions was obtained, as shown in Figure 12. From the figure, the crash risk of each unit under each weather condition can be clearly and visually identified, the influence of complex weather conditions on traffic crashes was highlighted, and the weather prone to crashes in each road section can be identified.

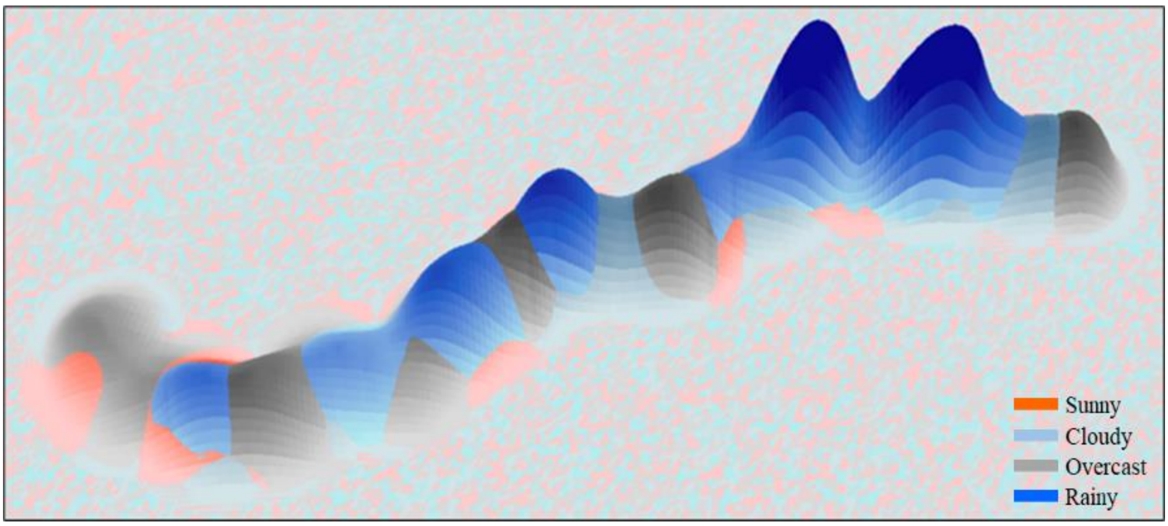


Figure 12. Crash risk distribution in complex weather.

The effective search index verifies the identification effect, as shown in Table 3.

Table 3. Identification effect of the time-spatial density ratio method.

Weather	Risk Level	P_{ts}	Ratio to the Total Length	Ratio to Total Crashes	Search Efficiency Index	
					Time-Spatial Density Ratio Method	Cumulative Frequency Method
Sunny	I II III	0–1.9 1.9–5.8 5.8–∞	0.818 0.158 0.025	0.139 0.592 0.269	0.17 3.749 10.994	5.466
Cloudy	I II III	0–1.9 1.9–5.8 5.8–∞	0.818 0.18 0.024	0.215 0.553 0.231	0.263 3.071 9.755	4.915
Overcast	I II III	0–1.9 1.9–5.8 5.8–∞	0.789 0.185 0.026	0.18 0.513 0.307	0.229 2.777 11.588	4.553
Rainy	I II III	0–1.9 1.9–5.8 5.8–∞	0.749 0.213 0.039	0.081 0.576 0.342	0.109 2.706 8.892	4.826

According to the calculation results in Table 3, the proportions of road lengths for the three risk levels in each weather condition are shown in Figure 13.

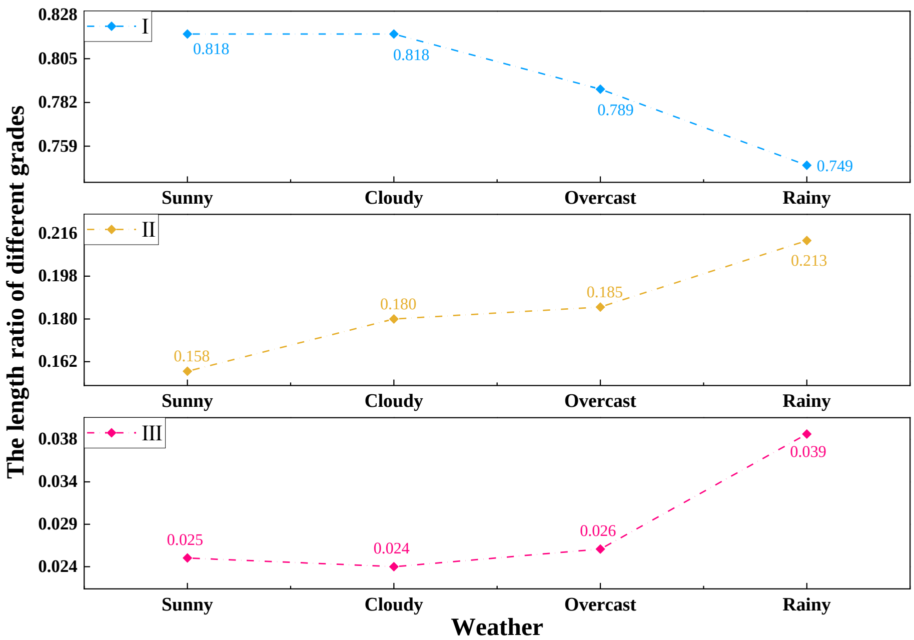


Figure 13. The length ratio of each weather condition and each grade.

There is a clear trend in Figure 13. In level I, the proportion of sunny and cloudy days was relatively large, and the proportion of overcast and rainy days was relatively reduced. In level II, the proportion of sunny days was relatively small, while the proportion of other weather gradually increased. In level III, the proportion of rainy days increased significantly. It can be seen that the weather has a significant impact on the classification of crash risk levels.

5. Comparative Analysis and Discussion

In the case of roads that cannot be identified uniformly due to differences in the road network, the time-spatial density ratio method and the cumulative frequency method are based on the relative risk level of road units. Crash-prone sections are determined according to the threshold by sorting them. This allows for macro-control of the road's crash tendency. The time-spatial density ratio method has the following advantages compared with the cumulative frequency method. By considering the crash distribution under complex weather conditions and comprehensive parallel ranking, the crash-prone sections under different weather conditions were identified. By utilizing the threshold determination method suitable for the characteristics of the curve in this study, the accuracy of threshold and classification was improved. By quantifying the impact of road length and period on crash distribution, the accuracy of road risk identification was improved.

However, the following explanation is required on the calculation method. On the one hand, the road cross section and the number of lanes in the research section have not changed. The highway road environment is similar. So the road section units were not distinguished from the perspective of road attributes. On the other hand, the period of the research data was extensive. It was not easy to quantify the traffic volume in units of days. More importantly, the primary purpose of this study was to determine the crash-prone sections of the entire road more accurately. Therefore, factors other than crash distribution were not considered.

In this case, when identifying crash-prone sections from a spatial perspective, the identification results covered 36.0% of the crashes in 6.4% of the road. The effective search index was increased by 89.9% compared with the cumulative frequency method. From a time-spatial perspective, on sunny days, 2.5% of the road covered 26.9% of sunny crashes, and the effective search index increased by 101.1%; on cloudy days, 2.4% of the road covered 23.1% of cloudy crashes, and the effective search index increased by 98.5%; on overcast days, 2.6% of the road covered 30.7% of overcast crashes, and the effective search index increased by 154.5%; on rainy days, 3.9% of the road covered 34.2% of rainy crashes, and the effective search index increased by 84.3%. According to the comparison of the effective search index, the identification effect of the time-spatial density ratio method was better in each situation. According to the calculation results, the crash-prone sections on rainy days were much greater than in other weathers. It can be seen that the time-spatial density ratio method has weather pertinence and high accuracy. Based on statistical theory, this method does not require established model assumptions and has the characteristics of simple calculation.

Another advantage of the time-spatial density ratio method is that it can be applied in many ways, providing a reference for road management. Through spatial analysis, the crash-prone sections in spatial can be quickly determined, indicating which sections need to be fixed and fitted out with safety features. The impact of different weather conditions on the unit or the road can be identified through time analysis, and the weather-specific safety decision-making and safety facility layout plan can be put forward. Through time-spatial analysis, the crash risk situation of highways under various weather conditions can be comprehensively presented, which offers a foundation for the traffic management department to control the traffic in real-time according to weather and road sections. Furthermore, hazard ranking helps tackle the priority problem that typically emerges in highway operation management due to a lack of people and financial resources.

Based on the results of this study, the locations of crash-prone sections under different weather conditions are different. The impact of weather on traffic safety should be actively considered. Meteorological monitoring, data sharing, and dynamic early warning should be the primary means of safety management and control. The management department should prepare emergency prompts, crash prevention, and measures in crash-prone sections according to weather conditions.

6. Conclusions

Identifying crash-prone sections is critical to the planning of traffic policies and the implementation of safety measures. Crash-prone sections are identified using a variety of methods. Considering the differences of highways and the practical applicability, the cumulative frequency method identifies crash-prone sections based on the number of crashes and the relative risk of road sections. However, previous studies have pointed out that the weather environment significantly impacts driving safety. In order to consider the weather factor and improve the identification accuracy, this paper proposed the time-spatial density ratio method, which conducted a targeted study on the safety of mountain highway sections under complex weather conditions.

Crashes were linked to weather conditions, road length, and period using the time-spatial density ratio method. The crash distribution was calibrated by distinguishing between different weather conditions. In this paper, combined with the actual data analysis of the G76 Expressway in the southwestern mountainous area of China, the identification accuracy was more than 80% higher than that of the cumulative frequency method. The crash-prone sections under different weather conditions were found, revealing the weather variances.

The distribution of road crashes in mountainous areas is influenced by a number of factors. Considering the difficulty in obtaining detailed crash information and the inaccuracy of crash data on many highways, this study only used the number of crashes as the basis for risk analysis. To continuously enhance identification accuracy, elements impacting crash distribution such as crash severity, road property damage, and traffic flow can be considered based on richer data.

In summary, this study analyzed the impact of complex weather on the crash risk of highways in mountainous areas. Given the current data situation in traffic research, the time-spatial density ratio method is an effective tool to obtain crash-prone sections under complex weather conditions accurately. Through the introduction and application of this paper, we hope to encourage researchers to explore the use of this highly applicable statistical method.

At present, the historical data of many roads are not detailed enough due to the different methods of determining traffic statistics. New statistical methods are needed to enable researchers to extract more precise results from historical data. These easy-to-apply methods certainly have general applicability and meteorological pertinence to the data of different roads. This type of statistical analysis holds considerable promise in the absence of data detail. As research into statistical analysis of highway crash data progresses, researchers need to strengthen the underlying methodology. The pertinence, accuracy, and applicability of the methodology will be continuously improved in the future.

Author Contributions: Conceptualization, C.Z. and R.S.; methodology, R.S.; software, Y.X. and B.W.; validation, L.H., Y.X. and B.W.; formal analysis, Y.X.; investigation, R.S.; resources, R.S.; data curation, C.Z.; writing—original draft preparation, Y.X.; writing—review and editing, C.Z. and L.H.; visualization, B.W.; supervision, C.Z.; project administration, C.Z. and R.S.; funding acquisition, C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key Research & Development Program of China (grant number 2021YFB2600103, 2020YFC1512005, 2021YFB2301602), Sichuan Science and Technology Program (grant number 2022YFG0048), Sichuan Provincial Transportation Technology Project (grant number 2019-ZL-12, 2021-ZL-15, 2022-ZL-04).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data used in this manuscript are provided by government agencies, which are available for use in this study only.

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

References

1. Traffic Administration Bureau of the Ministry of Public Security. 2019 Statistical Bulletin on the Development of the Transportation Industry. 2020. Available online: http://www.gov.cn/xinwen/2020-05/12/content_5510817.htm (accessed on 13 April 2022).
2. WHO. Global Plan for the Decade of Action for Road Safety 2021–2030. 2021. Available online: https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/decade-of-action-for-road-safety-2021-2030 (accessed on 28 October 2021).
3. Yu, B.; Chen, Y.; Bao, S.; Xu, D. Quantifying drivers' visual perception to analyze accident-prone locations on two-lane mountain highways. *Accid. Anal. Prev.* **2018**, *119*, 122–130. [CrossRef] [PubMed]
4. Oppe, S. Development of traffic and traffic safety, global trends and incidental fluctuations. *Accid. Anal. Prev.* **1991**, *23*, 413–422. [CrossRef]

5. Zhang, J.; Wan, C.; He, A.; Zhang, D.; Soares, C.G. A two-stage black-spot identification model for inland waterway transportation. *Reliab. Eng. Syst. Saf.* **2021**, *213*, 107677. [CrossRef]
6. Prieto, F.; Gómez-Déniz, E.; Sarabia, J.M. Modelling road accident blackspots data with the discrete generalized Pareto distribution. *Accid. Anal. Prev.* **2014**, *71*, 38–49. [CrossRef] [PubMed]
7. Jordan, P. ITE and road safety audit-a partnership for traffic safety. *ITE J.* **1999**, *69*, 24–27.
8. Gregoriades, A.; Mouskos, K.C. Black spots identification through a Bayesian Networks quantification of accident risk index. *Transp. Res. Part C Emerg. Technol.* **2013**, *28*, 28–43. [CrossRef]
9. Shen, L.; Lu, J.; Long, M.; Chen, T. Identification of accident blackspots on rural roads using grid clustering and principal component clustering. *Math. Probl. Eng.* **2019**, *2019*, 2151284. [CrossRef]
10. Wan, Y.; He, W.; Zhou, J. Urban road accident black spot identification and classification approach: A novel grey verhuls–Empirical bayesian combination method. *Sustainability* **2021**, *13*, 11198. [CrossRef]
11. Yakar, F. A multicriteria decision making–based methodology to identify accident-prone road sections. *J. Transp. Saf. Secur.* **2021**, *13*, 143–157. [CrossRef]
12. Zhang, C.; Shu, Y.; Yan, L. A Novel Identification Model for Road Traffic Accident Black Spots: A Case Study in Ningbo, China. *IEEE Access* **2019**, *7*, 140197–140205. [CrossRef]
13. Wang, J.; He, S.; Zhai, X.; Wang, Z.; Fu, X. Estimating mountainous freeway crash rate: Application of a spatial model with three-dimensional (3D) alignment parameters. *Accid. Anal. Prev.* **2022**, *170*, 106634. [CrossRef] [PubMed]
14. Saccomanno, F.F.; Grossi, R.; Greco, D.; Mehmood, A. Identifying black spots along highway SS107 in Southern Italy using two models. *J. Transp. Eng.* **2001**, *127*, 515–522. [CrossRef]
15. Malyshkina, N.V.; Mannering, F.L. Zero-state Markov switching count-data models: An empirical assessment. *Accid. Anal. Prev.* **2010**, *42*, 122–130. [CrossRef] [PubMed]
16. Barua, S.; El-Basyouny, K.; Islam, M.T. Effects of spatial correlation in random parameters collision count-data models. *Anal. Methods Accid. Res.* **2015**, *5*, 28–42. [CrossRef]
17. Debrabant, B.; Halekoh, U.; Bonat, W.H.; Hansen, D.L.; Hjelmborg, J.; Lauritsen, J. Identifying traffic accident black spots with Poisson-Tweedie models. *Accid. Anal. Prev.* **2018**, *111*, 147–154. [CrossRef]
18. Chang, L.Y.; Chen, W.C. Data mining of tree-based models to analyze freeway accident frequency. *J. Saf. Res.* **2005**, *36*, 365–375. [CrossRef]
19. Zhang, Y.; Lu, H.; Qu, W. Geographical detection of traffic accidents spatial stratified heterogeneity and influence factors. *Int. J. Environ. Res. Public Health* **2020**, *17*, 572. [CrossRef]
20. Fan, Z.; Liu, C.; Cai, D.; Yue, S. Research on black spot identification of safety in urban traffic accidents based on machine learning method. *Saf. Sci.* **2019**, *118*, 607–616. [CrossRef]
21. Gutierrez-Osorio, C.; Pedraza, C. Modern data sources and techniques for analysis and forecast of road accidents: A review. *J. Traffic Transp. Eng.* **2020**, *7*, 432–446. [CrossRef]
22. Ahmed, M.; Huang, H.; Abdel-Aty, M.; Guevara, B. Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway. *Accid. Anal. Prev.* **2011**, *43*, 1581–1589. [CrossRef]
23. Yu, R.; Abdel-Aty, M.; Ahmed, M. Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors. *Accid. Anal. Prev.* **2013**, *50*, 371–376. [CrossRef] [PubMed]
24. Yakar, F. Identification of accident-prone road sections by using relative frequency method. *Promet Traffic Transp.* **2015**, *27*, 539–547. [CrossRef]
25. 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] [PubMed]
26. Kwon, O.H.; Park, M.J.; Yeo, H.; Chung, K. Evaluating the performance of network screening methods for detecting high collision concentration locations on highways. *Accid. Anal. Prev.* **2013**, *51*, 141–149. [CrossRef] [PubMed]
27. Wei, C.; Wen, H. Geochemical baselines of heavy metals in the sediments of two large freshwater lakes in China: Implications for contamination character and history. *Environ. Geochem. Health* **2012**, *34*, 737–748. [CrossRef]
28. Fan, K.; Wei, C.Y.; Yang, X.S. Geochemical baseline of heavy metals in the soils of Qiaokou Town, Changsha City and its application. *Acta Sci. Circumstantiae* **2014**, *34*, 3076–3083.
29. Johnson, B.A.; Scheyvens, H.; Khalily, M.B.; Onishi, A. Investigating the relationships between climate hazards and spatial accessibility to microfinance using geographically-weighted regression. *Int. J. Disaster Risk Reduct.* **2019**, *33*, 122–130. [CrossRef]
30. Xie, Z.; Yan, J. Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: An integrated approach. *J. Transp. Geogr.* **2013**, *31*, 64–71. [CrossRef]
31. Anderson, T.K. Kernel density estimation and K-means clustering to profile road accident hotspots. *Accid. Anal. Prev.* **2009**, *41*, 359–364. [CrossRef]
32. Mohaymany, A.S.; Shahri, M.; Mirbagheri, B. GIS-based method for detecting high-crash-risk road segments using network kernel density estimation. *Geo Spat. Inf. Sci.* **2013**, *16*, 113–119. [CrossRef]
33. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*, 1st ed.; Routledge: New York, NY, USA, 1998; pp. 1–176. [CrossRef]
34. Wu, P.; Meng, X.; Song, L. Identification and spatiotemporal evolution analysis of high-risk crash spots in urban roads at the microzone-level: Using the space-time cube method. *J. Transp. Saf. Secur.* **2021**, *14*, 1510–1530. [CrossRef]