WILEY | Hindawi

Research Article

Research on Operation Characteristics and Safety Risk Forecast of Bus Driven by Multisource Forewarning Data

Shejun Deng (ID)**, Hongru Yu** (ID)**, and Caoye Lu** (ID)

College of Civil Science and Engineering, Yangzhou University, YangZhou, JiangSu 225009, China

Correspondence should be addressed to Shejun Deng; yzrx6@163.com

To prevent and control public transport safety accidents in advance and guide the safety management and decision-making optimization of public transport vehicles, based on the forewarning and other multisource data of public transport vehicles in Zhenjiang, holographic portraits of public transport safety operation characteristics are constructed from the perspectives of time, space, and driver factors, and a prediction model of fatigue driving and driving risk of bus drivers based on BP neural network is constructed. Finally, model checking and virtual simulation experiments are carried out. The result of the research shows that the driver's fatigue risk during the period of 7 : 00-8 : 00 am is much higher than other periods. When the bus speed is about 30 km/h, the driver fatigue forewarning events occur the most. Drivers aged 30–34 years have the largest proportion of vehicle abnormal forewarning, drivers aged 40–44 years have the largest proportion of fatigue forewarning events, and drivers with a driving experience of 15–19 years have the largest overall proportion of various forewarning events. When the vehicle speed range is (18, 20) km/h and (42, 45) km/h, the probability of fatigue driving risk and driving risk forewarning increases sharply; and when the vehicle speed is lower than 17 km/h or 41 km/h, the probability of fatigue driving risk and driving risk forewarning, respectively, is almost zero. The probability of fatigue forewarning during low peak hours on rainy days is about 30% lower than that during peak hours. The probability of driving forewarning during flat peak hours is 15% higher than that during low peak hours and about 10% lower than that during peak hours. This paper realized for the first time the use of real forewarning data of buses in the full time, the whole region, and full cycle to carry out research. Related results have important theoretical value and practical significance for scientifically guiding the safety operation and emergency management strategies of buses, improving the service level of bus passenger transportation capacity and safety operation, and promoting the safety, health, and sustainable development of the public transportation industry.

1. Background Introduction

At present, China mainly evaluates the safety of buses based on the incidence of traffic accidents. The evaluation indicators and analysis methods are relatively single, and there is still a lack of accurate control, effective prevention, and emergency management countermeasures. Since 2019, with the integration and system development of BDS, video, radar, and other technologies, buses in some Chinese cities have installed vehicle driving safety forewarning systems, enabling holographic perception, dynamic monitoring, and risk reminders of the bus operation process [1]. By acquiring the historical data of vehicle forewarning of Zhenjiang Public Transport Company in Jiangsu Province of China,

this paper realized for the first time the use of real forewarning data of buses in the full time, the whole region, and full cycle to carry out research. This paper excavates the general rules and main hidden dangers of vehicle forewarning events and carries out objective analysis and situation prediction of bus operation risks [2]. Relevant research conclusions have important practical significance for improving the safe operation of buses, carrying out corresponding optimized dispatching [3] and emergency management, eliminating hidden dangers of bus operation, and improving and promoting the convenience, safety, and sustainable development of public transportation [4].

Many studies believe that the fatigued driver and driver driving state is the most important factor affecting urban

public transport safety, and the driver state is affected by the driver's attributes, external environment, and other aspects. Relevant scholars have researched related factors affecting the safe operation of buses, mainly as shown in Table 1. By studying the related factors that affect the severity of bus collisions [5], it can be seen that the factors such as start inhibition, automatic door opening, bus materials, and internal structure are relatively related to bus safety. Researches on perception and driving behavior [6] have shown that drivers who have experienced accidents are more likely to have collision accidents in the future. By studying the factors affecting road traffic accidents, it is known that the advanced driver assistance assessment system (ADAS) [7] can provide drivers with safety support and help avoid distractions. Besides, the vehicle anticollision forewarning strategy [8, 9] is formulated through the study of the driver's reaction time when a collision occurs. The Palm probability distribution method [10] is used to study road accident risk under different weather conditions. The research results show that the accident risk probability of snow is higher than that of rain. Among them, the greater the precipitation intensity, the higher the relative accident risk probability. Secondly, the logistic regression model [11] is used to study the correlation between the driver's age, gender, vehicle, road environment, and other factors and the severity of traffic accidents. The results show that the road infrastructure conditions and the driver's age have a significant impact on the severity of road traffic accidents. By using a logarithmic linear model [12] to study the impact of time factors on the severity of bus driver collision injuries, the results show that driving in the late night or early morning will increase the risk of serious injury to bus drivers.

Many scholars have researched vehicle safety characteristics and management requirements, mainly as follows. Firstly, utilizing the historical traffic data in the USA from 2005 to 2009, the potential risk factors of public transportation safety accidents are summarized [13], and it is found that the driver is the main factor in the occurrence of public transportation safety accidents. Studies have shown that as the driver's attention changes, there are significant differences in eye movement and gear operation [14]. At the same time, the characteristics of steering wheel operation and the characteristics of vehicle movement are also related to the characteristics of vehicle movement during lane changes [15]. Secondly, algorithms and models are used to analyze the causes and predictions of traffic accidents. By using the decision tree algorithm [16] to study the causes of vehicle collision accidents, the results show that human factors are the most important factor causing traffic accidents. According to the research results of the driver's steering characteristics, an evaluation model used to improve the steering stability of the car is established [17]. In [18], the backpropagation neural network model and generalized linear mixed model were used to analyze multi-source traffic data, which showed that flow plays an important role in vehicle collision prediction. Thirdly, a variety of models were built to better predict vehicle safety. A traffic accident model based on collaboration theory [19] was proposed to analyze accident scene data by combining

driving comfort thresholds. The dynamic prediction model of vehicle operation trajectory based on vehicle trajectory data [20] can calculate the suspicious collision position of the vehicle. The perceived safety of self-driving cars and their application value in transportation and road safety [21] were derived due to the analysis of the driving habits of 1,205 regular vehicle drivers. A hidden Markov model [22] is proposed by analyzing a large amount of traffic trajectory data, and it is verified that the model can better predict the occurrence of traffic conflicts.

Related scholars have also carried out many studies on the safe operation and management of buses. Firstly, preventive measures [23] are proposed through the identification and risk analysis of bus drivers' dangerous behaviors. Risk assessment and analysis of hazard sources of road traffic safety risks are carried out through the application of the road traffic safety risk index evaluation method [24], and a corresponding road traffic safety risk monitoring index system is constructed. Besides, aiming at the main problems of safety management, traffic safety management countermeasures [25] are proposed to reduce driver unsafe behavior, improve vehicle safety level, reduce the accident rate, and ensure the safe operation of buses. Secondly, the safety of the driver's visual perception of dangerous areas is proposed by analyzing the eye movement data of changing lanes, cornering driving, and straight driving [26]. An evaluation method based on the driver's visual perception of safety indicators is established. In [27], a psychological fatigue evaluation system for bus drivers was constructed, and the authors proposed targeted suggestions to reduce driving fatigue. A method [28] that can evaluate the driver's potential danger prediction ability and the rationality of the system was designed. In [29], a method to analyze the safety operation of buses was proposed based on big trajectory data. In particular, research on the clustering characteristics of road safety factors [30] such as driver, vehicle, road, and environment under different accident types was conducted. According to the actual situation of vehicle safety prediction, different research methods are proposed to make the prediction results more accurate. Thirdly, in the establishment of the bus speed model, parameters such as bus flow and bus ratio [31] were introduced, and a bus speed control system was designed, which realized the dynamic monitoring of the vehicle running speed. The driver evaluation system based on the principal component analysis method [32] was established through the analysis and investigation of the questionnaire information of bus drivers. The research results show that the driver's driving habits and individual characteristics have a significant impact on driving behavior. Finally, by introducing the practice of traffic congestion charging in Singapore and London [33], it is concluded that public transportation congestion charging should be based on scientific planning and supporting the sustainable development of public transportation. Besides, the analysis method and test method of the index system [34] are used to classify the sustainable development of urban transportation. The evaluation index system and evaluation model of urban transportation sustainable development based on the theory of urban transportation sustainable

TABLE 1: Study of the relevant factors affecting the safe operation of buses.

Author	Year	Research
Gibson and Lee	1986	Driver's reaction time when a collision occurs
Mercier et al.	1999	Correlation between the driver's age, gender, vehicle, road environment, and the severity of traffic accidents
Chen et al.	2000	Influence of time factors on the severity of collision injuries of bus drivers
Yang	2007	Perception and driving behavior
M. Staubach	2009	Factors affecting road traffic accidents
Cafiso	2013	Related factors affecting the severity of bus collisions
Park	2019	Road accident risk under different weather conditions

development are established. This research provides a new perspective on urban sustainable development research.

In summary, objective, real, comprehensive, and effective historical operation data are a prerequisite for the accurate study of bus safety operation situation and risk management. The existing research studies mainly use vehicle accident data, vehicle trajectory data, laboratory data, and questionnaire survey data to carry out related research on vehicle safety characteristics, dangerous driving behavior, or risk situation. Because of the contingency of vehicle accidents and the incompleteness of data collection, it is difficult to realize the comprehensive analysis of bus operation state and the accurate prediction of safety risks [35]. This paper will overcome the shortcomings of the existing research, make full use of the safety forewarning system installed on public vehicles, obtain the real mass historical data of all kinds of public transport forewarning, carry out model construction and simulation analysis, provide auxiliary decision-making for bus operation, dispatching, and safety management, and promote the healthy, green, and sustainable development of urban public transportation [36].

2. Data Acquisition Process

The bus forewarning system installed by Zhenjiang Public Transport Company integrates various technologies such as ADAS yaw forewarning [37], fatigue driving video analysis, and BDS terminal [38]. It can realize the real-time upload of vehicle operating data and ensure the accuracy and reliability of the data. The forewarning equipment is shown in Figure 1. This research makes full use of the vehicle forewarning equipment and vehicle forewarning data platform of Zhenjiang Public Transport Company to obtain historical operating data of the bus.

2.1. Forewarning Equipment

2.1.1. ADAS Vehicle Yaw Forewarning System. ADAS stands for the Advanced Driving Assistance System. The system uses a camera located on the windshield to monitor the lane markings on the road ahead. When the system detects that the vehicle has deviated from the lane [39], it will issue a forewarning to the driver.

2.1.2. Fatigue Driving Analysis Equipment. The fatigue driving analysis equipment uses advanced AI video analysis

technology [40] to accurately recognize the driver's facial characteristics. At the same time, it can record and warn the driver's fatigue characteristics.

2.1.3. BDS. The artificial satellite's multifrequency positioning signal can be accepted by the system to achieve precise positioning. The system can calculate the distance of the vehicle ahead, consider the relative speed of the vehicle, determine the possible collision time, and issue a forewarning to the driver.

2.2. Forewarning System and Data Platform. The forewarning system can realize real-time monitoring and summary of vehicle information [41], mainly including 7 forewarning types: eyes closed, yawn, glance about, lane departure, rapid acceleration, rapid deceleration, and forward collision. This paper classifies the types of forewarnings, classifies eyes closed, yawn, and glance as driver fatigue forewarnings, and classifies rapid acceleration, rapid deceleration, forward collision, and lane departure as vehicle abnormal state forewarnings, as shown in Figure 2.

Data platforms mainly include current online, forward forewarning, driver forewarning, the total number of abnormalities, vehicle distribution, forewarning type distribution, forewarning occurrence trend, and other data. This paper obtained 297,189 forewarning data from November 2019 to March 2020 through the forewarning platform system of Zhenjiang Public Transport Company. The original forewarning data mainly include information such as license plate number, forewarning time, forewarning type, forewarning level, forewarning speed, latitude and longitude coordinates of forewarning points, location of forewarning points, driver names, and other information.

2.3. Research Period and Weather Conditions. Since the system started trial operation at Zhenjiang Public Transport Company in October 2019, this paper selected a 27-day official operation period from November 2019 to March 2020 to conduct research, as shown in Table 2.

2.4. Data Cleaning. Since the actual data obtained may have data missing, disordered format, abnormal data, and other phenomena, cleaning the data is an indispensable link. The principle of data cleaning [42] should ensure the accuracy, completeness, consistency, uniqueness, timeliness, and

(a)   (b)   (c)
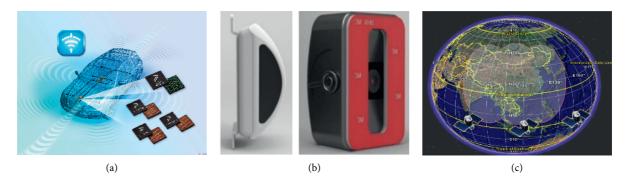
Figure 1: Forewarning equipment of buses: (a) ADAS yaw forewarning system; (b) fatigue driving analysis equipment; (c) BDS.
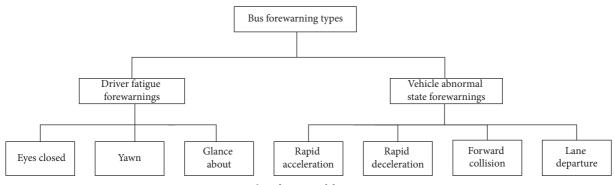

Figure 2: Classification of forewarning types.

Table 2: Weather and period of the retrieval date.

Date	Weather	Period
2019.11.01	Sunny	06 : 00–20 : 30
2019.11.02	Foggy	00 : 00–24 : 00
2019.12.01	Rainy	10 : 00–22 : 30
2019.12.02	Sunny	05 : 30–19 : 00
2019.12.04	Sunny	06 : 00–23 : 30
2019.12.05	Sunny	00 : 00–24 : 00
2019.12.06	Sunny	00 : 00–24 : 00
2019.12.07	Sunny	01 : 00–24 : 00
2019.12.15	Sunny	07 : 30–19 : 00
2019.12.16	Sunny	06 : 30–18 : 00
2019.12.17	Rainy	00 : 30–24 : 00
2019.12.18	Rainy	00 : 00–24 : 00
2019.12.20	Sunny	12 : 30–18 : 30
2019.12.21	Rainy	01 : 00–24 : 00
2020.01.01	Snowy	12 : 30–18 : 30
2020.01.02	Sunny	07 : 30–15 : 30
2020.01.03	Rainy	11 : 00–15 : 00
2020.01.04	Rainy	07 : 30–11 : 00
2020.01.06	Rainy	16 : 30–19 : 30
2020.01.07	Rainy	10 : 30–18 : 00
2020.01.08	Rainy	09 : 00–23 : 00
2020.01.09	Snowy	01 : 00–24 : 00
2020.01.10	Cloudy	00 : 00–24 : 00
2020.01.11	Rainy	12 : 30–19 : 00
2020.01.12	Rainy	08 : 30–23 : 30
2020.01.13	Foggy	00 : 00–24 : 00
2020.01.14	Foggy	00 : 00–24 : 00

effectiveness of data. There are mainly the following 4 methods [43]:

(1) Supplement incomplete data

(2) Detection and resolution of error values or abnormal values

(3) Detection and elimination of duplicate records

(4) Detection and resolution of data inconsistencies

The forewarning data set obtained and the driver information data set are associated and fused with the data table [44]. Finally, 297189 forewarning samples are associated with 1435 driver data samples. The research data samples after the fusion are shown in Table 3.

3. Analysis of the Bus Forewarning Characteristics from Multiple Perspectives

There are many factors related to the bus safe operation, and different factors have different effects on bus forewarning [45, 46]. To study the bus forewarning characteristics in different weather conditions, sections, driver characteristics, and periods, this paper makes a holographic portrait of the bus operation from multiple perspectives such as weather, time, space, speed, and driver characteristics distribution [47], as shown in Figure 3. Make full use of various multisource forewarning data to study the influence mechanism of various factors on bus forewarning.

TABLE 3: Data sample after cleaning.

Bus number	Time	Type	Longitude	Latitude	Route	Type number	Speed
18200148	5 : 38 : 33	Eyes closed	119.46	32.16	No. 81 bus	8	20
1709270416	8 : 18 : 53	Smoking	119.47	32.19	No. 35 bus	12	34
1709270762	8 : 19 : 54	Smoking	119.44	32.14	No. 95 bus	12	43
1709270416	8 : 21 : 29	Smoking	119.46	32.19	No. 35 bus	12	41
1709270041	8 : 21 : 58	Smoking	119.42	32.21	No. 23 bus	12	42
1709270762	8 : 22 : 11	Smoking	119.44	32.13	No. 95 bus	12	50

FIGURE 3: Portrait analysis of the safe operation of buses.

3.1. Weather Distribution.

Vehicle operation safe is closely related to bad weather conditions [48]. This paper compares and analyzes the forewarning data of buses according to the four weather conditions of sunny, rain, fog, and snow. The results are shown in Table 4.

From the analysis of Figure 4, it can be seen that the proportion of forewarnings on sunny and foggy days is relatively large. The total proportion of forewarnings on sunny days reaches 31.29%. It is mainly due to the glare of the sun and dizziness. Drivers are easily sleepy and fatigued. The total proportion of forewarnings on foggy days reached 31.27%, mainly due to low air visibility, obstructed line of sight, and low road adhesion coefficient. Drivers need to maintain a high degree of attention for a long time and are prone to fatigue.

The vehicle's abnormal state forewarnings are greater than the number of driver fatigue forewarnings under all weather conditions. The vehicle's abnormal state forewarnings mainly refer to situations such as rapid acceleration, rapid deceleration, forward collision, and lane departure, which are not only closely related to the driver's bad driving behavior but also affected by road facilities, traffic environment, and other restrictive factors, resulting in a higher proportion.

3.2. Time Distribution.

According to the forewarning data of buses, statistics are summarized by time, and the result is shown in Figure 5. Analysis shows the following:

(1) Whether it is the forewarnings of driver fatigue state or the forewarnings of vehicle abnormal state, it shows a three-stage change rule of rising, local fluctuations, and falling over time.

(2) The first stage is a rapid rise period. At this time, the driver fatigue forewarning period is [4 : 00, 7 : 00], and the vehicle abnormal state forewarning period is [4 : 00, 9 : 00], which is 2 hours longer than the fatigue forewarning, mainly due to complicated traffic environment and other factors restrict.

(3) The second stage is a period of partial fluctuation. The driver fatigue forewarning and the vehicle abnormal state forewarning period are generally similar, within the range of [7 : 00, 19 : 00], and the vehicle abnormal state forewarning period is [9 : 00, 16 : 00].

(4) The third stage is the partial descent period, and the overall is divided into 2 changes. During the period of [16 : 00, 20 : 00], the reduction range is relatively large; during the period of [20 : 00, 22 : 00], the range of change is relatively small.

(5) The vehicle abnormal state forewarning is in [10 : 00, 11 : 00] and [16 : 00, 17 : 00] peaks, appearing in the two periods, and the driver fatigue forewarning is in the peak hours at [06 : 00, 09 : 00].

3.3. Spatial Distribution.

The driver fatigue forewarning and vehicle abnormal state forewarning data are imported into the electronic map of Zhenjiang, and the locations corresponding to the relevant forewarning samples are all mapped to the map, as shown in Figure 6.

Using the kernel density analysis method [49], the density distribution corresponding to the driver fatigue forewarning is obtained. From Figure 7, the following holds:

(1) The two forewarning types are generally consistent in the spatial distribution of urban road networks, but

TABLE 4: Distribution of bus forewarning types under different weather conditions.

Feature name	Feature coding	Total number		Driver fatigue		Vehicle abnormal state	
		Quantity	Percentage	Quantity	Percentage	Quantity	Percentage
Weather	Sunny	92994	31.29	40830	31.48	52164	31.14
	Rainy	85610	28.81	37852	29.19	47758	28.51
	Foggy	92936	31.27	40245	31.03	52691	31.46
	Snowy	25649	8.63	10766	8.3	14883	8.89

FIGURE 4: Distribution of forewarning types corresponding to different weathers.
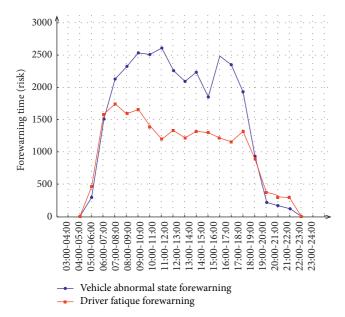

FIGURE 5: Time distribution characteristics of different forewarning types and number of buses.

there are large differences in local areas, as shown in Figure 7.

(2) The driver fatigue forewarning density center is concentrated on Zhongshan Road (Jiuhuashan Road-Jiefang Road), as shown in Figure 7(a). Zhongshan Road is the main road of Zhenjiang, and most of it is commercial land nearby. This area has a high density of people and traffic, which can easily cause vehicle congestion. Drivers need to maintain a high degree of tension for a long time in this complex traffic environment. It is prone to fatigue characteristics.

(3) The forewarning of vehicle abnormal state is mainly concentrated at the intersection of Dongwu Road and Mengxi Road, which is located in the Beigu Mountain Scenic Area, as shown in Figure 7(b). Beigu Mountain is an AAAAA-level tourist scenic spot in Zhenjiang, Jiangsu. The traffic volume of buses, private cars, tourist buses, and walking tourists is large, and the surrounding road network traffic congestion is serious. In this traffic environment, drivers are easy to make emergency operations and induce forewarning of abnormal vehicle status.

3.4. Speed Distribution. Select the period from December 2019 to January 2020 as the research period, analyze the corresponding vehicle speed when various forewarning events occur, obtain a total of 297189 data samples, and make summary statistics according to the speed, as shown in Table 5. Study the speed characteristic law under different forewarning types, as shown in Figure 8.

It can be seen from Figure 8 that as the speed increases, the number of driver fatigue forewarnings and the number of vehicle abnormal state forewarnings both fluctuate to a certain extent. The number of driver fatigue forewarnings reaches the peak at 30 km/h, and the number of driver forewarnings is less when the speed is less than 15 km/h or greater than 70 km/h. Since most buses operate in urban areas, the speed of running on urban roads is not high. When the speed of the vehicle is greater than 60 km/h, the number of driver fatigue forewarning and vehicle abnormal state forewarnings are kept at a low level. Both driver fatigue forewarning and vehicle abnormal state forewarnings have some abnormal values. The corresponding speed when the

(a)

(b)

Figure 6: Spatial distribution of forewarnings: (a) location of driver fatigue forewarning; (b) location of vehicle abnormal state forewarning.

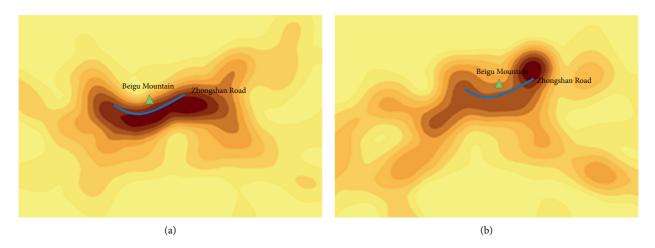(a)

(b)

Figure 7: Comparative analysis of partial forewarning areas: (a) driver fatigue forewarning; (b) vehicle abnormal state forewarning.

Table 5: Distribution of bus forewarning types under different speed conditions.

Feature name	Feature coding	Total number		Driver fatigue		Vehicle abnormal state	
		Quantity	Percentage	Quantity	Percentage	Quantity	Percentage
	Below 10 km/h	26225	8.82	7003	5.40	19222	12.96
	10–19 km/h	4307	1.45	511	0.39	3796	2.56
	20–29 km/h	46660	15.70	41571	32.05	5089	3.43
Speed	30–39 km/h	124333	41.84	42301	32.62	82032	55.32
	40–49 km/h	65313	21.98	24469	18.87	40844	27.55
	50–59 km/h	23701	7.98	10172	7.84	13529	9.12
	60 km/h and above	6650	2.24	3666	2.83	2984	2.01

driver fatigue state forewarning occurs is generally greater than the speed when the vehicle abnormal state occurs.

To facilitate the analysis of the vehicle speed distribution characteristics under different forewarning density areas, this paper divides the forewarning occurrence areas into three types: low, medium, and high. The correlation analysis of the speed when the forewarning occurs in the three regions is carried out, and the characteristic law of the speed is studied, as shown in Figures 9 and 10.

It can be seen from Figure 9 that the low forewarning density area has the largest proportion when the speed is 30 km/h-39 km/h, and the smallest proportion when the speed is below 20 km/h. In areas with low forewarning density such as suburban areas, there are fewer vehicles, smooth roads, and a small number of forewarnings. The number of forewarnings is the highest when the speed of buses reaches about 35 km/h. The medium forewarning density area has the largest proportion when the speed is 30 km/h-39 km/h, and the smallest proportion when the speed is above 60 km/h; the high forewarning density area has the largest proportion when the speed is 20 km/h-29 km/h, and the smallest proportion when the speed is above 60 km/h. In high forewarning density areas such as the city center, due to traffic congestion, the speed of buses is slow, and a large number of forewarnings are generated. The number of forewarnings reaches a peak when the speed is
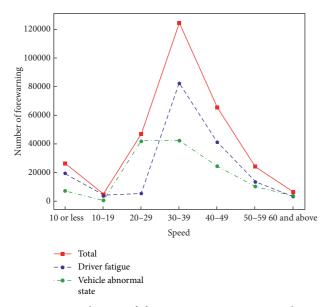
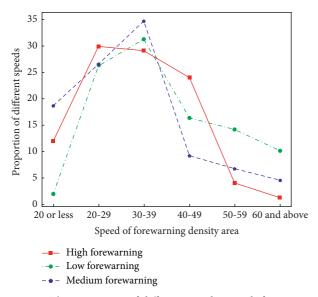FIGURE 8: Distribution of forewarning types corresponding to different speeds.

FIGURE 9: The proportion of different speeds in each forewarning density area.

about 25 km/h. Through observation and comparison, it can be seen that, due to different road congestion conditions, the speed of vehicles in high forewarning density areas is relatively low, and the speed of vehicles in low forewarning density areas is relatively high.

To be able to analyze the number of forewarnings that occur per unit area in each region more reasonably, this paper proposes the definition of unit forewarning density, that is,

$$\text{unit forewarning density} = \frac{\text{number of forewarnings}}{\text{forewarning area}}. \quad (1)$$

It can be seen from the analysis of Figure 10 that the forewarning frequency of each speed in the low warning
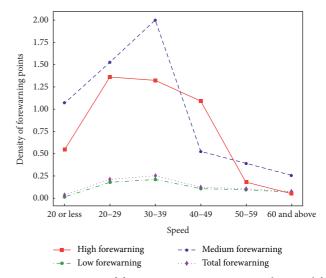

FIGURE 10: Density of forewarning points corresponding to different speeds in each forewarning density area.

density area is low. In the medium forewarning density area, the forewarning frequency reaches the peak when the speed is 30 km/h–39 km/h. In the high forewarning density area, when the speed is 20 km/h–29 km/h, the forewarning frequency is highest.

3.5. Driver Characteristics.

This paper takes 324 drivers of Zhenjiang Public Transport Company as the research object, analyzes the influence of drivers' age, driving years, gender, and educational background on the forewarning of buses [50], and conducts research on the distribution law of forewarning under the action of various factors.

3.5.1. Driving Years.

The statistical analysis of the forewarning data of bus drivers of different driving years is shown in Table 6; in terms of the total number of forewarnings, drivers with driving experience between 15 and 19 years of the four driving years have the most forewarnings, accounting for about 35.66% of the total forewarnings, and; drivers with a driving experience of fewer than 5 years have the least number of forewarnings, accounting for about 0.68% of the total number of forewarnings.

As shown in Figure 11, drivers with a driving experience of 15 to 19 years have the largest number of driver fatigue forewarnings and vehicle abnormal state forewarnings. Drivers in this age group are more daring after having certain driving experience, have more aggressive driving styles, and are prone to aggressive operations, so they are more prone to forewarning of abnormal vehicle conditions.

3.5.2. Age.

The statistical analysis of the forewarning data of bus drivers of different ages is shown in Table 7. In terms of the total number of forewarnings, drivers in the 40–44 age group of the eight age groups have the most forewarnings, accounting for about 23.13% of the total forewarnings. Drivers under the age of 25 have the least proportion, about 0.58%.

TABLE 6: Distribution of bus forewarning types for drivers of different driving years.

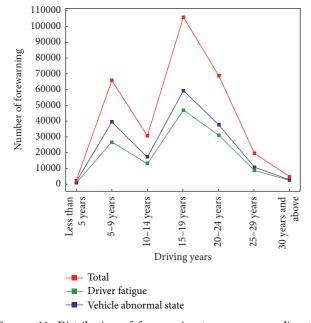| Feature name	Feature coding	Total number		Driver fatigue		Vehicle abnormal state	
		Quantity	Percentage	Quantity	Percentage	Quantity	Percentage
	Less than 5 years	2009	0.68	855	0.66	1154	0.69
	5–9 years	65889	22.17	26811	20.67	39078	23.33
	10–14 years	30452	10.25	13351	10.29	17101	10.21
Driving years	15–19 years	105964	35.66	46778	36.07	59186	35.34
	20–24 years	68677	23.11	31030	23.93	37647	22.48
	25–29 years	19281	6.49	8716	6.72	10565	6.31
	More than 30 years	4917	1.65	2152	1.66	2765	1.65

FIGURE 11: Distribution of forewarning types corresponding to drivers of different driving years.

As shown in Figure 12, from the perspective of different forewarning types, drivers in the 40–44 age group have the largest number of fatigue driving forewarnings, and drivers in the 30–34 age group have the largest number of vehicle abnormal state forewarnings. This is because 30–34 years old drivers have a more aggressive driving style, are more daring after having certain driving experience, and are easier to make aggressive operations. Drivers of different ages are generally more likely to have forewarnings of abnormal conditions than driver fatigue forewarnings. This is related to the fact that drivers are more aggressive in driving operations on the premise that they are safe.

3.5.3. Gender. Statistical analysis of the forewarning data of bus drivers of different genders is shown in Table 8; from the total number of forewarnings, male drivers have a much higher probability of having forewarnings than females. Among them, males account for approximately 91.44%, and the proportion of driver fatigue forewarning is about 91.26%, which has a certain relationship with the high proportion of male drivers in public transportation companies. As shown in Figure 13, comparing the number of warnings for male

and female drivers, it can be seen that the number of abnormal state forewarnings for drivers is higher than the number of fatigue forewarnings, while the numbers of fatigue forewarnings and abnormal state forewarnings for female drivers are very similar.

3.5.4. Degree. Statistical analysis of the forewarning data of bus drivers with different educational levels is shown in Table 9; drivers with junior high school and below have the most forewarnings, accounting for about 64.28% of the total number of forewarnings; drivers with high school education account for the least, about 12.98% of the total number of forewarnings. As shown in Figure 14, as the driver's educational background changes, the number of driver fatigue forewarnings and the number of vehicle abnormal state forewarnings show roughly the same changes.

4. Research on Risk Prediction of Public Transportation Safety Based on BP Neural Network Model

BP neural network is a concept proposed by Rumelhart and McClelland et al. It is a multilayer feedforward neural network with error backpropagation. The model has arbitrary complex pattern classification ability and excellent multibit function mapping ability and is suitable for complex nonlinear systems such as bus safety risk prediction.

4.1. Basic Principle

4.1.1. Structure of BP Neural Network. The topological structure of the BP neural network is shown in Figure 15, $\{x_1, x_2, \ldots, x_n\}$ is the input vector of BP neural network, $\{y_1, y_2, \ldots, y_n\}$ is the output vector of BP neural network, and ω_{ij} and ω_{jk} is the weight of BP neural network.

BP neural network can be regarded as a nonlinear function, and the network input value and predicted value are the independent variables and dependent variables of the function. When the number of input nodes is n and the number of output nodes is m, BP neural network expresses the functional mapping relationship from n independent variables to m dependent variables [51].

4.1.2. Error Backpropagation Algorithm. As a multilayer feedforward neural network, BP neural network is

TABLE 7: Distribution of forewarning types of buses for drivers of different ages.

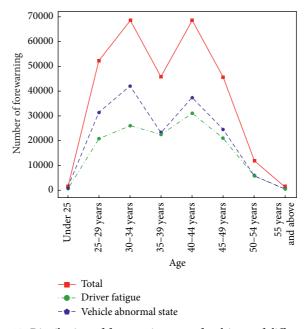| Feature name	Feature coding	Total number		Driver fatigue		Vehicle abnormal state	
		Quantity	Percentage	Quantity	Percentage	Quantity	Percentage
Age	Less than 25 years	1726	0.58	822	0.63	904	0.54
	25–29-year-old	52493	17.66	21007	16.20	31486	18.80
	30–34-year-old	68495	23.05	26302	20.28	42193	25.19
	35–39-year-old	46036	15.49	22649	17.46	23387	13.96
	40–44-year-old	68730	23.13	31137	24.01	37593	22.44
	45–49-year-old	45843	15.43	21206	16.35	24637	14.71
	50–54-year-old	12107	4.07	5831	4.50	6276	3.75
	55-year-old and above	1759	0.59	739	0.57	1020	0.61

FIGURE 12: Distribution of forewarning types for drivers of different ages.

TABLE 8: Distribution of forewarning types of buses for drivers of different genders.

Feature name	Feature coding	Total number		Driver fatigue		Vehicle abnormal state	
		Quantity	Percentage	Quantity	Percentage	Quantity	Percentage
Gender	Male	271454	91.34	118594	91.44	152860	91.26
	Female	25735	8.66	11099	8.56	14636	8.74

characterized by signal forward transmission and error backpropagation. In forward transmission, the input signal is processed layer by layer from the input layer through the hidden layer until the output layer. The neuronal states of each layer only affect the next layer of the neuron state. If the output layer cannot get the expected output, it will switch to backpropagation and continuously adjust the network weights according to the prediction error so that the predicted value of the model will converge gradually. The error backpropagation algorithm of the BP neural network [52] can be expressed as follows:

$$\delta^L = (t - y)f'\left(X^L - W^L\right),$$
$$\delta^l = \delta^{l+1}\left(W^{l+1}\right)^T f'\left(X^l W^l\right). \tag{2}$$

In formula (2), δ^l is the learning signal of layer l, δ^L is the learning signal of the output layer, t is the label value, y is the predicted value, X^l is the output signal of layer l, X^L is the output signal of the penultimate layer, W^l is the weight vector between the layer l and $l+1$, and W^L is the weight vector between the penultimate layer and the last layer.

The weight adjustment function of the BP neural network [53] is as follows:

$$\Delta W^L = -\eta \frac{\partial E}{\partial W^L},$$
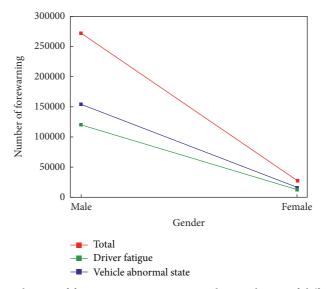$$\Delta W^l = -\eta \frac{\partial E}{\partial W^l}. \tag{3}$$

FIGURE 13: Distribution of forewarning types corresponding to drivers of different genders.

TABLE 9: Distribution of forewarning types of buses for drivers with different degree.

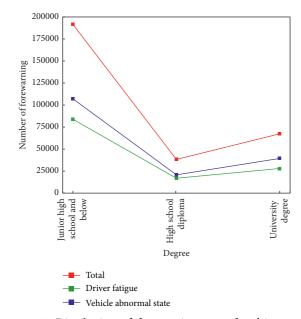| Feature name	Feature coding	Total number		Driver fatigue		Vehicle abnormal state	
		Quantity	Percentage	Quantity	Percentage	Quantity	Percentage
	Junior high school and below	191030	64.28	84208	64.93	106822	63.78
Education	High school diploma	67590	12.98	28030	21.61	39560	23.62
	University degree	38569	22.74	17455	13.46	21114	12.61

FIGURE 14: Distribution of forewarning types for drivers with different degrees.

In formula (3), ΔW^L is the adjustment value of the weight vector between the penultimate layer and the last layer of the BP neural network, ΔW^l is the adjusted value of the weight vector between the layer l and $l+1$, η is the learning rate, and E is the cost function.
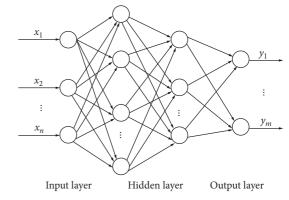

FIGURE 15: Topological structure diagram of BP neural network.

4.2. Model Building

4.2.1. Establish a Network Structure. Determining the network structure is an important part of constructing a BP neural network, which directly determines the training speed and prediction accuracy of the model. Generally speaking, the more hidden layers and nodes in the network topology structure, the stronger the generalization ability of the model, and the higher the accuracy of the model. However, the excessively complex network will lead to a slow training rate of the model and the more prone to overfitting; too simple network topology will make it difficult to establish a complex mapping relationship between feature variables and predictions, and it is difficult to achieve good

prediction results. Based on experience and repeated attempts, this paper confirms that the prediction results are good with the double hidden layer structure with node number of 100 and 50, respectively [54].

4.2.2. Selection of Learning Rate.

The learning rate is an important parameter in the process of model optimization, which determines the speed of model learning and the convergence effect of the model. Too much learning rate will cause the model accuracy to oscillate and be difficult to converge. Too small learning rate will lead to slow model adjustment. In this paper, 0.01 is selected as the learning rate of the model. At this time, the model converges faster and the oscillation amplitude is smaller [55].

4.2.3. Activation Function Selection.

The activation function in the BP neural network can increase the nonlinearity of the neural network so that the model has sufficient complex function mapping capabilities, and the applicability of different activation functions is also different [56].

(1) tanh Function. In this paper, the tan$h$ function is selected as the transfer function of the model. The tan$h$ function is the hyperbolic tangent function. It can maintain the non-linear monotonic rise and fall relationship on the output and input, which conforms to the gradient solution requirements of the BP network and has good fault tolerance and bounds. Besides, compared with the sigmoid activation function, tan$h$ function alleviates the problem of gradient disappearance to a certain extent, and its formula is as follows:

$$\tan h(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{4}$$

In formula (4), tan$h(x)$ is the function value of the hyperbolic tangent function, x is the input variable, and e is the natural constant.

(2) Softmax Function. In this paper, the softmax function is selected as the classifier of the model output. The softmax function is the normalized exponential function, which can normalize the gradient logarithm of the finite item discrete probability distribution. Its characteristic is to normalize the vector, highlight the maximum value, suppress other components far below the maximum value, and visually show that the sample is a certain type of confidence; the formula is as follows [57]:

$$\text{softmax}(X)_i = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}, \quad j = 1, 2, \ldots, n. \tag{5}$$

In formula (5), X is the input vector; softmax$(X)_i$ is the *i-th* function value of the vector softmax function for the vector X; x_i and x_j are the i and j values of the vector X, respectively; n the length of vector X; and the meaning of e is the same as above.

4.2.4. Cost Function Selection.

The cost function is mainly divided into two types: quadratic cost function and cross-entropy cost function [58]. The quadratic cost function is mainly used for regression problems. For the classification problems mentioned in this paper, the cross-entropy cost function is generally selected (labels are processed by one-hot encoding), and the formula is as follows:

$$E = -t(t \ln y + (1 - t)\ln(1 - y)). \tag{6}$$

In formula (6), E is the cost function value, t is the true label value, and y is the predicted value of the model.

Besides, the cross-entropy cost function also avoids the quadratic cost function: when the error is larger, the gradient of the activation function is smaller, resulting in slow convergence.

4.2.5. Data Preprocessing

(1) Normalization. To reduce the influence of the initialization value and accelerate the convergence speed of the BP neural network, the normalized preprocessing method can be generally adopted. In this paper, the maximum-minimum method is used to normalize the continuous characteristic variables [59], and the formula is as follows:

$$x_k = \frac{x_k - x_{\min}}{x_{\max} - x_{\min}}. \tag{7}$$

In formula (7), $x_{\min}$ is the minimum value of the feature in all samples, $x_{\max}$ is the maximum value of the feature in all samples, and x_k is the eigenvalue after normalization.

(2) One-Hot Encoding. To digitize classification and discrete variables into the model, it is necessary to map such features to Euclidean space. One-hot encoding is one of the most effective ways to achieve this function. One-hot encoding is also known as one-bit effective encoding and uses multibit status registers to encode multiple states: for a feature, if it has m values, it becomes m binary features after one-hot encoding.

4.3. Construction and Application of the Prediction Model.

According to the BP neural network model constructed in Section 4.2, 13 features such as weather conditions, driver data, driving period, and driving speed are taken as the input of the model, and the alarm of the driver is taken as the output of the model. The network topology structure of "13-100-50–2" is adopted, and the tan$h$ function and softmax function are used as the transfer function and activation function of the model, respectively. The cross-entropy cost function is selected as the cost function of the model, and after repeated attempts, the learning rate of the model is 0.01, which ensures the stable convergence of the model. The specific form of the model is shown in Figure 16 [60].

4.3.1. Fatigue Driving Prediction Model

(1) Investigation of Convergence and Dispersion. Randomly select 2/3 of the samples as the training set and the remaining 1/3 as the test set. Perform 500 cycles of iterative training on
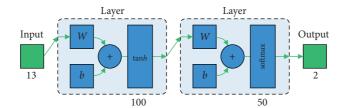
FIGURE 16: Safety state of buses model based on BP neural network.

the BP neural network. The learning curve of the fatigue driving prediction model is shown in Figure 17 [61].

As shown in Figure 17, the analysis shows the following: the fatigue driving prediction model has good convergence, and the learning curve tends to be flat around the 100th training cycle; during the whole 500-cycle iteration process, there was no large-scale oscillation and the fluctuation amplitude gradually decreased with the training cycle; the model performs well on the test set and can still reach an accuracy of 79% based on using a large number of static features; the model has no obvious overfitting in the training process, and there is only a 0.0034 accuracy difference between the set and the test set.

(2) Sample Inspection. Since the sample label adopts the form of one-hot encoding, with position 0 representing forewarning and position 1 representing no forewarning. Therefore, a single sample error can be obtained by randomly selecting the predicted value of the model with 200 samples and subtracting the true value, as shown in Figure 18.

It can be seen from Figure 18 that among the randomly selected prediction samples, the number of samples with correct forewarning accounts for 79%, 18% of the false positives are forewarnings, and only 3% of the samples are falsely reported as no forewarnings, which shows that the whole model is partial to safety and has high prediction accuracy under the application of state prediction.

4.3.2. Driving Risk Prediction Model

(1) Investigation of Convergence and Dispersion. Similar to the fatigue driving prediction model, 2/3 of the samples are randomly selected as the training set, and the remaining 1/3 are used as the test set. The BP neural network is trained iteratively for 300 cycles. The learning curve of the driving risk prediction model is shown in Figure 19.

From Figure 19, the following is obtained: the driving risk prediction model has good convergence. The learning curve tends to be flat around the 120th training cycle, but it fluctuates greatly from the 170th to the 210th cycle, which may be caused by the transformation of the model from the local optimal solution to the global optimal solution; The model reaches the highest state after 300 cycles of the iterative process, and the convergence speed was faster than that of fatigue driving risk forewarning model. The model is slightly weaker than the previous model in the test set, but it can still achieve a higher accuracy rate of 78%. The model has

no overfitting phenomenon in the training process, and the performance of the model in the test set is even better than that in the training set.

(2) Sample Inspection. Randomly select the predicted value of 200 samples from the model, and subtract the true value from it to get a single sample error, as shown in Figure 20.

As can be seen from Figure 20, among the randomly selected prediction samples, the number of samples with correct forewarning accounts for 78%, 14.5% of the false positives are forewarnings, and 7.5% of the samples are falsely reported as no forewarnings. The model is generally safe and has high prediction accuracy.

4.4. Research on Simulation of Risk Probability Prediction Based on the BP Model

4.4.1. Typical Driver Selection. This paper conducts a statistical analysis of 1565 drivers of the Zhenjiang Public Transport Company. For continuous features such as driving age and age, the mean value (16, 39) is used as the feature value of the virtual driver; for the classification features such as educational background and gender, the mode (high school, male) is taken as the characteristic value of typical drivers. An example of normalized virtual driver sample data is shown in Table 10.

4.4.2. Risk Probability Analysis during Peak Hours. The fatigue driving prediction model constructed in this paper is used to calculate the fatigue confidence of the virtual driver's sample data under different weather, periods, and speeds. The simulation results are shown in Figure 21.

It can be seen from the graph analysis that whether it is fatigue driving forewarning or driving risk forewarning, the probability of occurrence is positively increasing with the driving speed value; when the vehicle speed range is (18, 20) km/h and (42, 45) km/h, the probability of fatigue driving risk forewarning and driving risk forewarning, respectively, raises sharply; when the vehicle speed is lower than 17 km/h or 41 km/h, the probability of fatigue driving risk forewarning and driving risk forewarning, respectively, occurring is almost zero; under the same speed conditions, the probability of fatigue forewarning in snowy days is greater than that of foggy days, rainy days, and sunny days and the probability of driving forewarning in foggy days is greater than that of snowy days, rainy days, and sunny days.

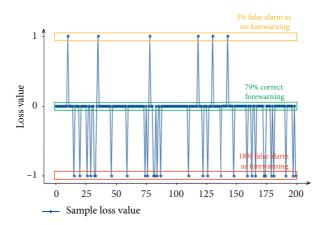Figure 17: Learning curve of fatigue driving prediction model.

Figure 18: Analysis diagram of sample loss value of fatigue forewarning prediction model.
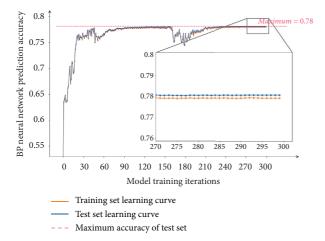

Figure 19: The learning curve of driving risk prediction model.

4.4.3. Risk Probability Analysis during Low Peak Hours. According to Figure 22, based on different speed conditions, the change characteristics of fatigue driving risk forewarning and driving risk forewarning probability are generally consistent with those in peak hours, indicating that high attention should still be paid to safe driving of vehicles in low peak hours; under the same speed conditions, the probability of fatigue forewarning in rainy days is about 30% lower than that in peak hours, and the difference in other weather conditions is small.

4.4.4. Risk Probability Analysis during Flat Peak Hours. According to the analysis in Figure 23, the change characteristics of fatigue driving risk forewarning and driving risk forewarning probability are generally consistent with peak hours and flat peak hours; under the same speed conditions, the probability of driving forewarning in four weather conditions is 15% higher than that in low peak hours and 10% lower than that in peak hours; at the same driving speed, the sequence of driving risk probability is foggy, snowy, rainy, sunny, and speed, indicating that driving risk is significantly related to weather conditions.
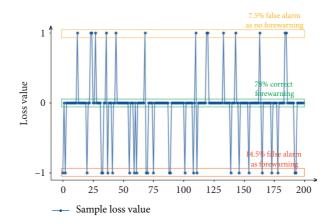
FIGURE 20: Analysis diagram of sample loss value of driving risk prediction model.

TABLE 10: Samples of typical drivers (partial).

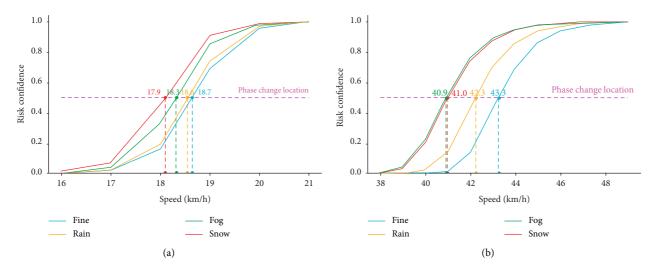|	Fine	Rain	Fog	Snow	PeakT	FlatT	LowT	Male	Female	Age	DrivingY	Edu	Speed
0	1	0	0	0	1	0	0	1	0	0.66	0.47	0.67	0.0051
1	1	0	0	0	1	0	0	1	0	0.66	0.47	0.67	0.0102
2	1	0	0	0	1	0	0	1	0	0.66	0.47	0.67	0.0152
80	1	0	0	0	0	1	0	1	0	0.66	0.47	0.67	0.0051
81	1	0	0	0	0	1	0	1	0	0.66	0.47	0.67	0.0102
160	1	0	0	0	0	0	1	1	0	0.66	0.47	0.67	0.0051
161	1	0	0	0	0	0	1	1	0	0.66	0.47	0.67	0.0102
240	0	1	0	0	1	0	0	1	0	0.66	0.47	0.67	0.0051
241	0	1	0	0	1	0	0	1	0	0.66	0.47	0.67	0.0102
242	0	1	0	0	1	0	0	1	0	0.66	0.47	0.67	0.0152
480	0	0	1	0	1	0	0	1	0	0.66	0.47	0.67	0.0051
481	0	0	1	0	1	0	0	1	0	0.66	0.47	0.67	0.0102
482	0	0	1	0	1	0	0	1	0	0.66	0.47	0.67	0.0152

(a)

(b)

FIGURE 21: (a) Comparison of fatigue driving risk. (b) Comparison of driving risk probability under different weather conditions.
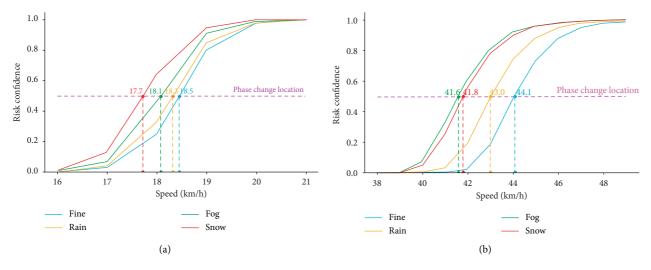
(a)

(b)

Figure 22: (a) Comparison of fatigue driving risk. (b) Comparison of driving risk probability under different weather conditions.

(a)

(b)

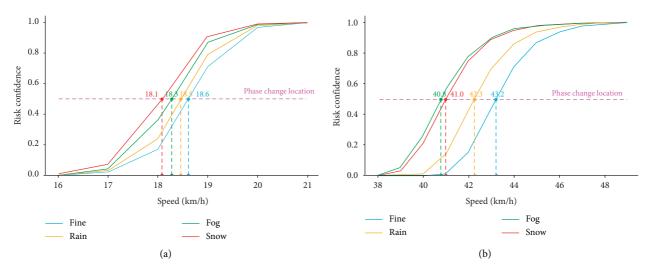Figure 23: (a) Comparison of fatigue driving risk. (b) Comparison of driving risk probability under different weather conditions.

5. Conclusions

5.1. Research Result. This paper selects 297189 various types of forewarning data of Zhenjiang buses to carry out the analysis of hidden risks and characteristic laws. The distribution characteristics of bus forewarnings of different weather conditions, speeds, periods, spaces, and driver characteristics are studied. We get the following conclusions: firstly, on sunny days from 7 : 00–8 : 00 am in the morning, the probability of driver fatigue forewarning is greatest. On foggy days from 11 : 00 am–12 : 00 noon, the probability of vehicle abnormal state forewarning is the greatest. Secondly, when the vehicle is running at 30 km/h, the proportion of driver fatigue forewarning is the largest. Urban core areas are prone to trigger forewarning of driver fatigue, while tourist attractions are prone to trigger vehicle abnormal forewarning. Finally, drivers with 15–19 years of driving experience have the largest proportion of fatigue forewarnings and vehicle abnormal forewarnings. Drivers aged 40–44 years have the largest proportion of fatigue forewarnings. Drivers aged 30–34 years have the largest proportion of

vehicle abnormal forewarnings; male drivers have the largest proportion of fatigue forewarnings and vehicle abnormal forewarnings.

The fatigue driving and driving risk prediction model based on BP neural network are constructed, and simulation analysis is performed. The results show that, at the same driving speed, the sequence of occurrence of driving risk probability is foggy, snowy, rainy, and sunny days. During peak hours, the probability of fatigue forewarning in snowy days is greater than that of foggy, rainy, and sunny days; the probability of driving forewarning in foggy days is greater than that of snowy, rainy, and sunny days. When the vehicle speed range is (18, 20) km/h and (42, 45) km/h, the probability of fatigue driving risk and driving risk forewarning increases sharply; when the vehicle speed is lower than 17 km/h or 41 km/h, the probability of fatigue driving risk and driving risk forewarning, respectively, is almost zero. The probability of fatigue forewarning during low peak hours on rainy days is about 30% lower than that during peak hours. The probability of driving forewarning during

flat peak hours is 15% higher than that during low peak hours and about 10% lower than that during peak hours.

5.2. Practical Implications. The relevant research conclusions of this paper are of great practical significance for improving the passenger transportation capacity of buses and enhancing the management level. At the same time, it can improve the auxiliary decision-making for the safe operation and emergency management of buses, promoting the sustainable and healthy development of urban public transport safety.

5.3. Limitation and Future Research Scope. This study was not free from limitations. Firstly, we selected 297189 samples from November 2019 to March 2020 with a total of 27 days, and the sample size is relatively small. Secondly, when studying the forewarning characteristics of different gender bus drivers, male drivers accounted for a larger proportion of the selected 324 drivers. Therefore, the conclusion that male drivers have a much higher forewarning rate than female drivers needs further verification. Finally, public transportation safety risk probability prediction is multifactor. Currently, it is only based on the actual data obtained by the forewarning equipment to predict.

Although the constructed model has certain accuracy, more influencing factors such as the type of road facilities, road traffic conditions, and types of bus stations need to be fully considered in the follow-up research.

Although the sample size is insufficient and there are some of the above shortcomings in the research, this paper realizes for the first time the use of real forewarning data of buses in the full time, the whole region, and full cycle to carry out research and the use of real data for objective evaluation, which is representative and innovative.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

S. D. and H. Y. conceptualized the study and prepared methodology. C. L. and H. Y. analyzed using software and investigated and visualized the study. S. D., C. L., and H. Y. validated the study. C. L. and S. D. performed data curation and wrote the original draft. S. D. supervised the study, administrated the project, did formal analysis, managed resources, obtained funding acquisition, and reviewed and edited the article.

Acknowledgments

References

[1] J. Tang, J. Liang, C. Han, Z. Li, and H. Huang, "Crash injury severity analysis using a two-layer stacking framework," *Accident Analysis and Prevention*, vol. 122, pp. 226–238, 2019.

[2] J. Tang, L. Zheng, C. Han et al., "Statistical and machine-learning methods for clearance time prediction of road incidents: a methodology review," *Analytic Methods in Accident Research*, vol. 27, pp. 1–16, Article ID 100123, 2020.

[3] Y. Wang, X. Ma, Z. Li, Y. Liu, M. Xu, and Y. Wang, "Profit distribution in collaborative multiple centers vehicle routing problem," *Journal of Cleaner Production*, vol. 144, pp. 203–219, 2017.

[4] Y. Yan, Y. Zhang, X. Yang, J. Hu, J. Tang, and Z. Guo, "Crash prediction based on random effect negative binomial model considering data heterogeneity," *Physica A*, vol. 547, pp. 1–12, Article ID 123858, 2020.

[5] C. Salvatore, D. G. Alessandro, and P. Giuseppina, "Road safety issues for bus transport management," *Accident Analysis and Prevention*, vol. 60, pp. 324–333, 2013.

[6] C. Y. D. Yang and Noblis, "Trends in transit bus accidents and promising collision countermeasures," *Journal of Public Transportation*, vol. 10, no. 3, pp. 077–291X, 2007.

[7] M. Staubach, "Factors correlated with traffic accidents as a basis for evaluating advanced driver assistance systems," *Accident Analysis & Prevention*, vol. 41, no. 5, pp. 1025–1033, 2009.

[8] J. J. Gibson, *The Ecological Approach to Visual Perception*, Rout ledge, Abingdon, UK, 1986.

[9] D. N. Lee, "A theory of visual control of braking based on information about time-to-collision," *Perception*, vol. 5, no. 4, pp. 437–459, 1976.

[10] S. J. Park, S. Y. Kho, S.-Y. Kho, and H.-C. Park, "The effects of road geometry on the injury severity of expressway traffic accident depending on weather conditions," *The Journal of the Korea Institute of Intelligent Transport Systems*, vol. 18, no. 2, pp. 12–28, 2019.

[11] C. R. Mercier, M. C. Shelley, G. H. Adkins, and J. M. Mercier, "Age and gender as predictors of injury severity in broadside and angle vehicular collisions," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1693, no. 1, pp. 50–61, 1999.

[12] W. Chen and P. P. Jovanis, "Method for identifying factors contributing to driver-injury severity in traffic crashes," *Transportation Research Record*, vol. 1717, no. 9, pp. 1–9, 2000.

[13] K. Sigal and G. P. Carlo, "Risk factors associated with bus accident severity in the United States: a generalized ordered logit model," *Journal of Safety Research*, vol. 43, pp. 171–180, 2012.

[14] F. Zong, H. Y. Zhang, and H. G. Xu, "Predicting severity and duration of road traffic accident," *Mathematical Problems in Engineering: Theory, Methods and Applications*, vol. 2013, Article ID 547904, 9 pages, 2013.

[15] N. Kuge, T. Yamamura, O. Shimoyama, and A. Liu, "A driver behavior recognition method based on a driver model framework," *SAE International Journal of Engines*, vol. 109, pp. 469–476, 2000.

[16] R. Vesna, B. Milan, and B. Branko, "Identifying the key risk factors of traffic accident injury severity on Slovenian roads using a non-parametric classification tree," *Transport*, vol. 32, no. 3, pp. 272–281, 2017.

[17] M. Abe and Y. Kano, "A study oil vehicle handling evaluation by model based driver steering behavior," in *Proceedings of the FISITA World Automotive Congress 2008, Congress Proceedings—Vehicle Concepts, Boay Design, Chassis Development*, pp. 377–384, Munich, Germany, June 2008.

[18] L. Mussone, M. Bassani, and P. Masci, "Analysis of factors affecting the severity of crashes in urban road intersections," *Accident Analysis & Prevention*, vol. 103, pp. 112–122, 2017.

[19] F. Chen, S. L. Chen, and Y. Y. Cheng, "Research on the model of road traffic safety evaluation based on synergetics theory," *Advanced Materials Research*, vol. 779-780, pp. 1028–1035, 2013.

[20] I. Md Tazul and E. B. Karim, "Full Bayesian evaluation of the safety effects of reducing the posted speed limit in urban residential area," *Accident Analysis and Prevention*, vol. 80, pp. 18–25, 2015.

[21] L. Montoro, S. A. Useche, F. Alonso, I. Lijarcio, P. Bosó-Seguí, and A. Martí-Belda, "Perceived safety and attributed value as predictors of the intention to use autonomous vehicles: a national study with Spanish drivers," *Safety Science*, vol. 120, pp. 865–876, 2019.

[22] K. Xie, K. Ozbay, H. Yang, and C. Li, "Mining automatically extracted vehicle trajectory data for proactive safety analytics," *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 61–72, 2019.

[23] C. Dai, *Research on the Post Characteristics of Modern Bus Drivers and Preventive Measures for Unsafe Behaviors*, Shandong University, Jinan, China, 2013.

[24] X. G. Zhao, *Research on Key Technologies of Dynamic Warning of Traffic Safety Risks in Regional Road Networks*, Chang'an University, Xi'an, China, 2010.

[25] M. Peng, *Research on Safety Management and Emergency Response of Urban Public Transport Operation*, Chang'an University, Xi'an, China, 2012.

[26] Y. Li, *Research on Multi-Attribute Evaluation Method and Application of Driving Behavior Safety*, Jilin University, Changchun, China, 2016.

[27] Y. Sang, *Research on the Mental Fatigue Evaluation of Beijing Public Transport Drivers*, Capital University of Economics and Business, Beijing, China, 2011.

[28] G. Q. Yan, *Evaluation of Drivers' Potential Hazard Predictive Ability*, Chang'an University, Xi'an, China, 2011.

[29] H. J. Ren, T. Xu, and X. Li, "Driving behavior analysis based on trajectory data collected with vehicle-mounted GPS receivers," *Geomatics and Information Science of Wuhan University*, vol. 39, no. 6, pp. 739–744, 2014.

[30] T. X. Sun, *Analysis of Road Traffic Accidents Based on Data Mining*, Beijing Jiaotong University, Beijing, China, 2014.

[31] C. W. Ren, *Research on Dynamic Supervision System of Bus Drivers*, Chang'an University, Xi'an, China, 2014.

[32] T. Yi, "Investigation and characteristic analysis of bad driving behavior of bus drivers——taking Chengdu as an example," *Heilongjiang Transportation Science and Technology*, vol. 260, no. 10, pp. 166–168, 2015.

[33] Y. Sun, "Research on urban road traffic congestion charging based on sustainable development," *Physics Procedia*, vol. 24, pp. 1567–1572, 2012.

[34] Y. Xing, H. Liang, and D. Xu, "Sustainable development evaluation of urban traffic system," *Procedia-Social and Behavioral Sciences*, vol. 96, pp. 496–504, 2013.

[35] Y. Yan, Y. Dai, X. Li, J. Tang, and Z. Guo, "Driving risk assessment using driving behavior data under continuous tunnel environment," *Traffic Injury Prevention*, vol. 20, no. 8, pp. 807–812, 2019.

[36] Y. Wang, S. Peng, X. Zhou, M. Mahmoudi, and Z. Lu, "Green logistics location-routing problem with eco-packages," *Transportation Research Part E: Logistics and Transportation Review*, vol. 143, Article ID 102118, 2020.

[37] Z. C. Guo, *Research on Lane Line Recognition and Lane Departure Judgment Method in ADAS*, Chang'an University, Xi'an, China, 2018.

[38] G. Chen and C. Zhao, "Design and implementation of vehicle monitoring platform based on Beidou high precision location," *Electronic Design Engineering*, vol. 28, no. 04, pp. 52–55, 2020.

[39] Y. Q. Xie, Z. B. Li, J. Qi, K. Zhang, B. X. Zhang, and F. Qi, "Optimal video communication strategy for intelligent video analysis in unmanned aerial vehicle applications," *Chinese Journal of Aeronautics*, vol. 33, no. 11, 2020.

[40] X. Q. Xue, *Research on Traffic Safety Early Warning Methods in the Environment of Internet of Vehicles*, Beijing Institute of Technology, Beijing, China, 2016.

[41] Y. Liang, K. L. Li, F. H. Bi, K. Zhang, and J. Yang, "Research on LFMCW radar velocity ranging optimization system based on FPGA," *Procedia Computer Science*, vol. 166, pp. 187–194, 2020.

[42] Y. F. Wang, C. Z. Zhang, B. B. Zhang, and T. T. Wu, "Overview of data cleaning research," *New Technology of Library and Information Service*, vol. 12, pp. 50–56, 2007.

[43] D. C. Corrales, A. Ledezma, and J. C. Corrales, "A case-based reasoning system for recommendation of data cleaning algorithms in classification and regression tasks," *Applied Soft Computing Journal*, vol. 90, 2020.

[44] H. J. Wu, F. Liu, L. Zhao, Y. B. Shao, and R. Cui, "Application research of crawler and data analysis based on Python," *International Journal of Advanced Network. Monitoring and Controls*, vol. 5, no. 2, pp. 64–70, 2020.

[45] H.-K. Chen, H.-W. Chou, J.-W. Su, and F.-H. Wen, "Structural interrelationships of safety climate, stress, inattention and aberrant driving behavior for bus drivers in Taiwan," *Transportation Research Part A: Policy and Practice*, vol. 130, pp. 118–133, 2019.

[46] H. L. Sun and L. Liu, "Main factors and countermeasures affecting the safety of urban public transportation in my country," *Biotech World*, vol. 10, no. 10, p. 44, 2012.

[47] B.-r. Cao, H.-H. Cao, and Y. Liu, "Research on construction and application of an evaluation system for regional road traffic safety," *Grey Systems: Theory and Application*, vol. 8, no. 4, pp. 527–538, 2018.

[48] A. Theofilatos, "Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials," *Journal of Safety Research*, vol. 61, pp. 9–21, 2017.

[49] Y. Ge and K. Li, "Research on layout planning of automobile charging pile based on nuclear density analysis and space syntax——a case of Wuhan in hubei," *Urbanism and Architecture*, vol. 16, no. 34, pp. 41–45, 2019.

[50] V. Petraki, A. Ziakopoulos, and G. Yannis, "Combined impact of road and traffic characteristic on driver behavior using

smartphone sensor data," *Accident Analysis and Prevention*, vol. 144, 2020.

[51] J. Y. Sun and Y. F. Su, "Application of BP neural network based on matlab in public transport scheduling," *Journal of Wuhan University of Technology (Transportation Science & Engineering)*, vol. 35, no. 3, pp. 599–602, 2011.

[52] Y. F. Zhang, S. Y. Kuang, J. Liang, and J. Xu, "A recognition method for multi-reconnaissance data based on error back propagation algorithm," *Electronic Warfare Technology*, vol. 34, no. 6, pp. 18–22, 2019.

[53] P. Chen, *Research on Intelligent Real-Time Dispatching Model and System Implementation of Public Transport Based on BP Neural Network*, Beijing Jiaotong University, Beijing, China, 2008.

[54] L. Wang and Y. J. Kuang, "Research of BP neural network optimizing method based on ant colony algorithm," *New Industrialization Straregy*, vol. 2, no. 4, pp. 8–15, 2012.

[55] J. Li, C. P. Chen, H. M. Deng, and Y. Ma, "Short-term traffic flow prediction based on improved variable learning rate BP algorithm," *Automation & Instrumentation*, vol. 1237, no. 2, pp. 182–184, 2016.

[56] Y. Xuan and R. L. Wang, "Research on the improvement of dual optimization on BP neural network," *Journal of Guangxi Teachers Education University (Natural Science Edition)*, vol. 35, no. 1, pp. 60–65, 2018.

[57] Y. D. Zhang, "Simulation of real-time management of orderly access information for urban traffic in big data," *Computer Simulation*, vol. 35, no. 12, pp. 139–142, 2018.

[58] L. Chen, W. He, and L. Q. Han, "Radial basis function neural network modeling of the traffic path cost function," *CAAI Transactions on Intelligent Systems*, vol. 06, no. 5, pp. 424–431, 2011.

[59] L. Yu, "Rolls-like algorithm of self-driving cars——whether the principle of "maximum minimum" can be used as the moral decision-making principle of "tramcar problem"" *Philosophical Trends*, vol. 10, pp. 100–107, 2019.

[60] C. J. Cai, "Research on short-term traffic flow forecast model of BP neural network," *Fujian Computer*, vol. 3, pp. 14-15, 2015.

[61] W. Liu, H. X. Liu, and T. Zhu, "BP neural network for vehicle state estimation was based on ROC curve optimization," *Journal of Xi'an University of Architecture & Technology*, vol. 46, no. 4, pp. 593–597, 2014.