Innovative Methods for Data Informed Multimodal Transport Flow and System Analysis 2020

Lead Guest Editor: Zhiyuan Liu
Guest Editors: Weiwei Qi, Ye Yu, and Yanyong Guo

Innovative Methods for Data Informed Multimodal Transport Flow and System Analysis 2020

Innovative Methods for Data Informed Multimodal Transport Flow and System Analysis 2020

Lead Guest Editor: Zhiyuan Liu
Guest Editors: Weiwei Qi, Ye Yu, and Yanyong Guo

Hongtai Yang (iD), China
Vincent F. Yu (iD), Taiwan
Mustafa Zeybek, Turkey
Jing Zhao, China
Ming Zhong (iD), China
Yajie Zou (iD), China

Contents

WILEY | Hindawi

Research Article

Analyzing the Safety Impacts of Variable Speed Limit Control on Aggregated Driving Behavior Based on Traffic Big Data

Xu Qu [ID],[1,2] **Mofeng Yang** [ID],[3] **Junyi Ji** [ID],[1] **Linheng Li** [ID],[1,2] **and Bin Ran** [ID][1,2,4]

[1]*School of Transportation, Southeast University, Nanjing, China*
[2]*Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University Road #2, Nanjing 211189, China*
[3]*Department of Civil and Environmental Engineering, University of Maryland-College Park, College Park, MD 20742, USA*
[4]*Department of Civil and Environmental Engineering, University of Wisconsin-Madison, 1204 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706, USA*

Correspondence should be addressed to Xu Qu; quxu@seu.edu.cn

Variable speed limit (VSL) control dynamically adjusts the displayed speed limit to harmonize traffic speed, prevent congestions, and reduce crash risks based on prevailing traffic stream and weather conditions. Previous research studies examine the impacts of VSL control on reducing corridor-level crash risks and improving bottleneck throughput. However, less attention focuses on utilizing real-world data to see how compliant the drivers are under different VSL values and how the aggregated driving behavior changes. This study aims to fill the gap. With the high-resolution lane-by-lane traffic big data collected from a European motorway, this study performs statistical analysis to measure the difference in driving behavior under different VSL values and analyze the safety impacts of VSL controls on aggregate driving behaviors (mean speed, average speed difference, and the percentage of small space headway). The data analytics show that VSL control can effectively decrease the mean speed, the speed difference, and the percentage of small space headways. The safety impacts of VSL control on aggregated driving behavior are also discussed. The aggregated driving behavior variables follow a trend of first decreasing and then increasing with the continuous decrease in VSL values, indicating that potential traffic safety benefits can be achieved by adopting suitable VSL values that match with prevailing traffic conditions.

1. Introduction

Variable speed limit (VSL) control dynamically adjusts the displayed speed limit on the variable message signs to harmonize traffic speed, prevent congestions, and reduce crash risks based on prevailing traffic stream and weather conditions, which is an essential control strategy for Active Traffic Management (ATM) system. Germany was among the earliest countries that implemented the VSL system in the mid-1960s [1]. Nowadays, VSL control has been widely adopted by countries including Germany, America, the Netherlands, United Kingdom, and China for real-time operation of motorway traffic to improve traffic safety and mitigate congestions [2].

To evaluate the safety benefits of the VSL control, many approaches have been taken. Questionnaire surveys are used to evaluate the safety benefits of VSL [3, 4]. Their results showed that around 95% of drivers believe that VSL can effectively improve driving safety. The driving simulator is another way to examine the impact of VSL on driving behavior. A previous study shows that VSL can improve traffic safety by reducing speed differences between vehicles [5]. However, both questionnaire surveys and driving simulators are biased towards respondents' subjective factors, which might lead to the overestimation of VSL systems' safety benefits. Traffic simulation methods have been widely used to examine both the operation and safety benefits of different ATM control strategies [6–9]. To better reproduce the real-

world effect of VSL control, different traffic simulation methods have also been applied to evaluated VSL benefits by simulating different driver groups' behavioral responses to the VSL [7–9]. Some researchers integrated a real-time crash potential prediction model that utilizes the output of the microscopic simulation [10–13]. They found that VSL could achieve safety benefits, especially in reducing the risk of rear-end collisions. The simulation-based studies are able to quantitatively evaluate VSL safety impacts; however, the driving behavior is heterogeneous between different driver groups. For instance, aggressive drivers tend to drive with a higher desired speed compared to conservative drivers [14]. The microscopic simulation models simply apply the compliance rate to the driver group and universally set the compliant drivers' desired speed equalling to VSL values, which had not been proven correct by any empirical observations.

Empirical data, e.g., the loop detector data, has been widely used to analyze different aspects of motorway traffic [15, 16]. With the support of these high-resolution data, driving behaviors under VSL control have also been well studied [17–23]. The consensus emerging from these studies is that VSL control can improve safety by decreasing the mean speed, the speed difference, and the percentage of small-time headway and increase the mean time headway. Studies using empirical traffic data as inputs can objectively analyze the impacts of VSL on driving behaviors and derive credible conclusions. However, with the constraint of pre-defined VSL control strategies, the majority of these studies only have access to data from a limited range of traffic conditions and limited VSL values. Therefore, most of these studies merely performed qualitative comparisons between driving behavior parameters with and without VSL controls, while missing the quantitative examination of the relationship between traffic variables and VSL values. In the real-world implementation, VSL values vary in accordance with the prevailing traffic conditions and weather. Therefore, such studies were not able to evaluate the potential outcomes of each VSL value with a limited amount of data. In the meantime, few studies have controlled traffic states when examining VSL impacts; therefore, it is difficult to determine whether the obtained results are caused by VSL control or by the difference in traffic states within the collected empirical data.

According to the literature, the understanding of VSL control's mechanisms leading to efficiency and safety is still not fully explored and understood by researchers. Efforts devoted to investigating the impacts of VSL on driving behaviors have been insufficient, especially for quantitative research studies on VSL. To the best knowledge of the authors, few have examined the impact of VSL control on the aggregated driving behaviors with different VSL values [24].

To fill the gap, this study utilized high-resolution lane-by-lane traffic big data from a European motorway under different VSL values. Then a broad range of traffic states, i.e., traffic density, is categorized and investigated under various VSL values. With controlling the traffic state, the relationships between aggregated driving behaviors and VSL values

within different traffic density intervals have been quantified. Based on the quantitative analysis results, this study measures the difference in driving behavior under different VSL values and discusses the safety impacts of VSL control on aggregated driving behaviors and potential improvement on the motorway rear-end collisions. The results of this study revealed the mechanism leading to the safety benefits of VSL control and provided more realistic assumptions for modeling traffic flow operations under VSL control. Furthermore, we also discussed the invaluable insights into developing effective VSL control strategies to improve traffic safety.

2. Analyzing Aggregated Driving Behavior with Traffic Big Data

2.1. Traffic Big Data. The high-resolution lane-by-lane traffic big data is collected from a two-direction European motorway segment with three lanes in each direction. Inductive loop detectors are placed every 500-meter on the motorway collecting traffic stream speed, headway, vehicle length, and other traffic flow characteristics. The motorway segment experiences a wide variety of traffic conditions (including recurrent and nonrecurrent congestions). A VSL control system is deployed along the motorway with the control objective of alleviating roadway congestions and improving traffic safety. A wide variety of speed limits including 50 km/h, 60 km/h, 80 km/h, 100 km/h, and 120 km/h are implemented in the system.

This study uses traffic flow data and corresponding VSL control speeds data from 7:00 am to 6:00 pm for two weeks at a no weaving area site. The data only in good weather condition (no rain or fog) and visibility conditions (at daytime) is selected to avoid external interference. Missing and wrong data was removed. Besides, in order to eliminate drivers' adjustments immediately after the VSL control is triggered, the traffic flow data within 3 minutes after the implementation of VSL control is removed. Finally, the dataset including 4266 minutes data records with 355,599 vehicles was established, including 128,998 vehicles, 136,301 vehicles, and 90,300 vehicles on the left lane, middle lane, and right lane, respectively.

2.2. Traffic States. The investigation of the VSL impacts on driving behavior must control the traffic states. Under various traffic states, the driving behavior will be differently affected by the VSL control system. Therefore, the classification of the samples according to traffic states is indeed critical to this study. In this study, traffic density is chosen as a critical indicator to measure traffic congestion. The samples are further classified into different traffic density intervals to evaluate the impact of VSL control under respective traffic states. And the samples in the same density interval with and without VSL controls are also compared to investigate the effect of VSL control on driving behavior.

Furthermore, traffic data needs to be aggregated for the density parameter of every sample. The traffic volume, the mean speed, and the density of every sample are calculated

using one minute as the statistical time interval. The traffic flow data within different density intervals under different VSL control speeds on different lanes are classified by five vehicles per kilometer per lane density intervals. Meanwhile, to ensure there are enough samples to perform statistical analysis, the density intervals under each VSL value whose sample size is too low should be ignored. Through trial and error, fifteen is selected as the lower limit of the sample size. The number of samples within different density intervals under different speed limits on different lanes is listed in Table 1 ("—" indicates that the number of samples is less than 15 within the associated density interval). Samples with high traffic density (larger than 30 veh/km) are removed because, under high traffic density, the reduction of traffic speed is mainly caused by traffic congestions and stop-and-go waves, thus making it hard to quantify the actual impact of VSL on the aggregated driving behavior.

2.3. Aggregated Driving Behavior Variables.

In this study, we defined three variables to represent the aggregated driving behavior: mean speed, average speed difference, and the percentage of small space headway. These three variables are calculated using the traffic big data and are further used to measure the safety impact of the VSL control. Definition and calculation for these three aggregated driving behavior variables are shown below.

2.3.1. Mean Speed.

The two most frequently used mean speeds in traffic studies are the time mean speed and the space mean speed. This study selected the space mean speed as one of the analysis variables. It is the arithmetic mean speed value of vehicles within a certain section at a moment, which equals the harmonic mean of spot speed observed.

$$\overline{v}_s = \frac{n}{\sum_{m=1}^{n} 1/v_m}, \tag{1}$$

where v_m represents the instantaneous velocities of the vehicle m; n represents the number of vehicles passing the roadway segment.

2.3.2. Average Speed Difference (ASD).

The average speed difference of two neighboring vehicles was selected to measure the speed difference. It is the mean value of speed difference between the neighboring vehicles passing a given point in a specified time interval [25].

$$\text{ASD} = \overline{\Delta v} = \frac{\sum_{i=1}^{n-1} |v_i - v_{i+1}|}{n-1}, \tag{2}$$

where v_i represents the speed of the ith vehicle passing the fixed point; n represents the number of vehicles passing a given point.

2.3.3. The Percentage of Small Space Headway.

Space headway is a measurement of the minimum possible distance between vehicles without a reduction in the speed of the following vehicles. A small disturbance might still lead to instability conditions on motorways when a platoon of vehicles is operating with small headways [26]. Therefore, the distribution of headway, particularly the percentage of small headways, has some influence on the stability and safety of traffic stream operations. The percentage of less than 100 meters of space headway was used as one of the analysis variables.

3. Results

The impact of VSL control on aggregated driving behavior is measured based on the mean speed, the variation of speeds, and headways in this study as introduced before.

3.1. VSL Impacts on Mean Speed.

In order to compare the effect of different VSL values on the mean speed, the arithmetic average of all the sample's mean speed within different density intervals is calculated under different speed limits on different lanes. The results are summarized in Table 2. The two-sample Student's t-test is used to compare the mean speeds' difference with and without VSL controls, and the associated p values for t-tests are also provided in parenthesis to infer the impact significance of the VSL control. The null hypothesis (H_0) is that the index is the same for the conditions with and without control. If the t-test value is less than 0.05, we accept the hypothesis, if not, we reject it. The t-statistics is calculated using the formula shown below:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}, \tag{3}$$

where $\overline{x}_1$, $\overline{x}_2$ respectively, represent sample means of the two groups, s_1^2, s_2^2 represent the sample variances, and n_1, n_2 represent the sample sizes.

The results in Table 2 reveal that the mean speeds are lower under VSL controls regardless of the traffic density interval or lane locations. Under noncongested traffic density conditions (5–20 veh/km), the speed limit of 120 km/h only provides a marginal reduction in mean speeds on the left lane and the middle lane, yet the t-test results show that these speed differences are still statistically significant; meanwhile, other lower VSL speed values (lower than 120 km/h) are significantly reducing the mean speeds, and the t-test results for these speeds are also statistically significant. Under congested conditions (25–30 veh/km on the left and the middle lane, 20–25 veh/km on the right lane), the t-test results reveal that the speed limit of 120 km/h is not significantly changing the mean traffic stream speed when compared to conditions without VSL control; for other lower VSL values, obvious mean speed reductions are achieved, and these mean speed differences are all statistically significant.

To evaluate the impact of different VSL values for each lane under the same traffic state, enough samples need to be obtained under each VSL value for a given traffic density interval. Given the sample size information shown in Table 1, only the traffic density interval 20–25 veh/km has enough sample size under all VSL values. Therefore,

TABLE 1: Number of samples within different density intervals under different speed limits ("—" means the sample size is less than 15).

Lane	Density (veh/km)	Number of samples under different speed limits (km/h)					
		No VSL	120 km/h	100 km/h	80 km/h	60 km/h	50 km/h
Left lane	0–5	217	28	18	—	—	—
	5–10	621	147	83	—	—	—
	10–15	496	178	103	18	—	—
	15–20	317	395	94	30	16	—
	20–25	221	687	54	48	45	15
	25–30	101	315	19	—	—	—
Middle lane	5–10	97	21	—	—	—	—
	10–15	476	194	79	—	—	—
	15–20	1020	804	219	60	15	—
	20–25	345	637	66	18	30	16
	25–30	50	119	—	—	—	—
Right lane	5–10	325	175	31	—	—	—
	10–15	1220	875	242	49	—	—
	15–20	412	626	99	31	38	15
	20–25	37	91	—	—	—	—

TABLE 2: The arithmetic average and t-test results of mean speed within different density intervals under different speed limits on different lanes ("—" means the sample size is less than 15).

Lane	Density (veh/km)	The arithmetic average of mean speed under different speed limits (km/h)					
		No VSL	120 km/h	100 km/h	80 km/h	60 km/h	50 km/h
Left lane	0–5	129.6	125.7(0.00)	111.4(0.00)	—	—	—
	5–10	126.1	124.8(0.00)	109.9(0.00)	—	—	—
	10–15	122.5	121.4(0.00)	106.3(0.00)	98.0(0.00)	—	—
	15–20	117.6	113.2(0.00)	102.1(0.00)	91.3(0.00)	91.4(0.00)	—
	20–25	111.6	106.3(0.00)	96.4(0.00)	86.5(0.00)	86.4(0.00)	90.6(0.00)
	25–30	100.6	99.6(**0.56**)	85.2(0.00)	—	—	—
Middle lane	5–10	116.9	110.2(0.00)	—	—	—	—
	10–15	115.1	111.3(0.00)	98.5(0.00)	—	—	—
	15–20	111.8	104.3(0.00)	96.7(0.00)	84.4(0.00)	82.0(0.00)	—
	20–25	105.1	97.8(0.00)	91.8(0.00)	84.0(0.00)	80.6(0.00)	87.0(0.00)
	25–30	85.7	88.9(**0.18**)	—	—	—	—
Right lane	5–10	98.8	92.5(0.00)	88.8(0.00)	—	—	—
	10–15	97.6	90.7(0.00)	88.6(0.00)	78.4(0.00)	—	—
	15–20	92.6	85.9(0.00)	86.0(0.00)	78.1(0.00)	72.7(0.00)	77.0(0.00)
	20–25	80.4	81.3(**0.67**)	—	—	—	—

20–25 veh/km was selected for the left lane and middle lane, and 15–20 veh/km was selected for the right lane to demonstrate the change trends in the mean speed under different speed limits. As shown in Figure 1, it can be observed that, after the 60 km/h speed limit, the mean speeds increase with the speed limits reduction. The reason might be that some drivers might challenge the rationality of VSL controls when a low-speed limit is implemented, whereby the reduced VSL compliance rate leads to the phenomenon of higher mean speeds under lower VSL values.

3.2. VSL Impacts on ASD. In order to compare the effects of different speed limits on the ASD, the arithmetic average of all the sample's ASD within different density intervals is calculated under different speed limits on different lanes. Table 3 summarizes the results. Again, Student's t-test is used to compare the ASDs' difference between VSL control

groups and No VSL groups, and the associated p values are also provided in parenthesis.

It can be observed from Table 3 that, for all the lanes within the same density intervals, the ASDs are always lower under VSL controls. In addition, the t-test results show that the difference in ASDs is statistically significant for the majority of conditions except for speed limits of 120 km/h on the left and middle lane within low-density intervals and speed limits of 60 km/h on the left lane within the 15–20 veh/km density interval.

Figure 2 describes the change trends in ASD under different speed limits on the left and middle lane within the 20–25 veh/km density interval and on the right lane within the 15–20 veh/km density interval. It is clear that ASDs first decrease and then increase with the reduction in speed limits. A possible explanation for this phenomenon is that when VSL speed values are low, those aggressive drivers might choose to retain their speeds or only adopt a small
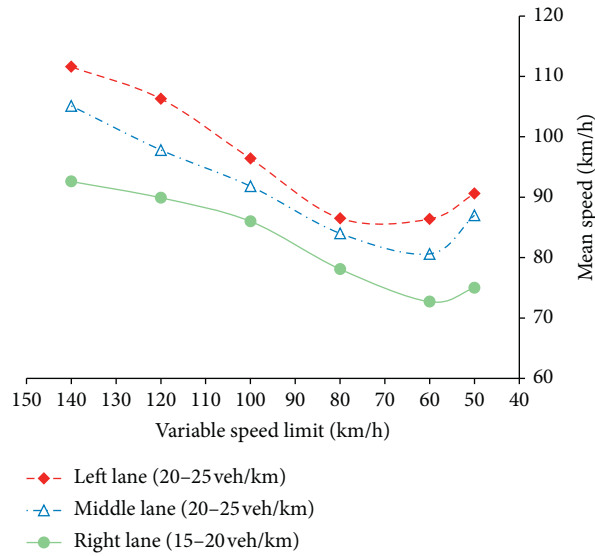
FIGURE 1: Change trends in mean speed under different speed limits.

TABLE 3: The arithmetic average and *t*-test results of ASD within different density intervals under different speed limits on different lanes ("—" means the sample size is less than 15).

Lane	Density (veh/km)	The arithmetic average of ASD under different speed limits (km/h)					
		No VSL	120 km/h	100 km/h	80 km/h	60 km/h	50 km/h
Left lane	0–5	8.89	7.01(0.01)	6.85(0.02)	—	—	—
	5–10	6.54	6.23(**0.08**)	5.88(0.00)	—	—	—
	10–15	5.13	4.97(**0.14**)	4.53(0.00)	4.41(0.03)	—	—
	15–20	4.20	3.90(0.00)	3.53(0.00)	3.62(0.00)	4.11(**0.73**)	—
	20–25	3.61	3.24(0.00)	3.12(0.00)	2.95(0.00)	3.15(0.00)	3.19(0.03)
	25–30	3.10	2.87(0.00)	2.90(0.03)	—	—	—
Middle lane	5–10	7.80	7.78(**0.97**)	—	—	—	—
	10–15	6.35	5.92(0.00)	4.55(0.00)	—	—	—
	15–20	5.37	4.69(0.00)	3.82(0.00)	3.44(0.00)	5.03(0.00)	—
	20–25	4.43	3.62(0.00)	3.19(0.00)	3.38(0.00)	3.26(0.00)	3.67(0.01)
	25–30	3.32	2.86(0.00)	—	—	—	—
Right lane	5–10	10.19	8.75(0.01)	5.63(0.00)	—	—	—
	10–15	7.91	6.33(0.00)	4.57(0.00)	3.47(0.00)	—	—
	15–20	6.02	4.48(0.00)	3.55(0.00)	3.10(0.00)	3.64(0.00)	4.15(0.01)
	20–25	4.21	3.58(0.01)	—	—	—	—

speed reduction, while those conservative drivers will follow the speed limits and apply a large speed reduction. Such behavioral differences lead to a more discrete distribution for drivers' desired speeds, whereby resulting in higher ASD values in the traffic flow.

3.3. VSL Impacts on Space Headway.

In order to quantify the effects of different VSL values on the space headway, the percentage of small headways within different density intervals is calculated under different speed limits on different lanes. Table 4 summarizes the percentage of vehicles with smaller than 100 m space headway under all combinations of

density intervals and speed limits on the left, the middle, and the right lane ("—" indicates that less than 300 vehicles are observed within the associated density interval).

It can be observed from the above table that the majority of percentage (of less than 100 m headway) results on the left and middle lane are smaller for VSL control groups when compared to No VSL groups. However, most percentage results on the right lane with VSL control are similar to the results of No VSL controls except for the speed limit of 100 km/h. The speed limit of 60 km/h on the right lane increases the percentage of less than 100 m headway. In general, the percentages of less than 100 m headway first decrease and then increase for all lane locations within the
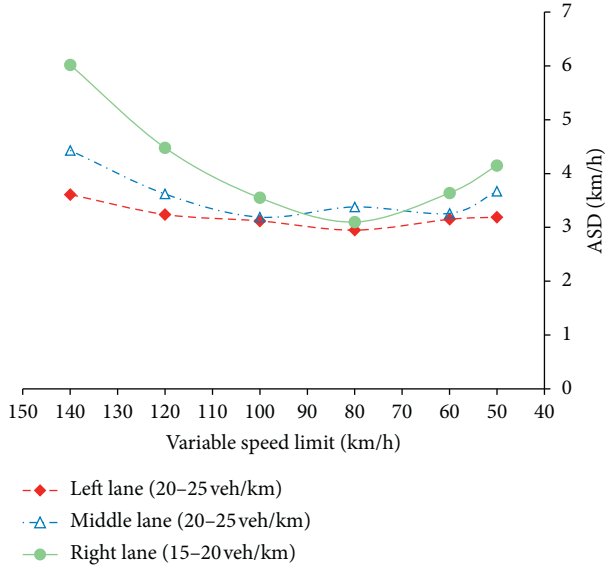
FIGURE 2: Change trends in ASD under different speed limits.

same density intervals with the lowest rate achieved at the speed limit of 100 km/h. Further reduction in speed limits beyond 100 km/h will not further reduce the percentage of small-time headways.

3.4. VSL Impacts on Motorway Rear-End Collisions. This section focuses on understanding the impacts of VSL on rear-end collisions because rear-end collision is the most frequent accident type on the motorway. Although some rear-end collisions have resulted from inappropriate driver behaviors, vehicle failure, and so on, most collisions are closely related to the vehicle speed, the car following distance, and the speed difference between consecutive vehicles.

It is assumed that two vehicles are running in the same lane on a motorway. As shown in Figure 3, the leading vehicle is running at the speed of V_1, and the following one at the speed of V_2 with the distance of d from the former. If an emergency happened ahead forcing the leading vehicle driver to brake immediately and adjust its speed to V_0 at the deceleration rate of a, the following driver also needs to brake to avoid a collision and needs to adjust its speed to at most V_0. The following vehicle's deceleration is assumed to be the same as the leading vehicle, and the drivers' reaction times are ignored.

The distance of the leading vehicle traveled:

$$S_1 = \frac{V_0^2 - V_1^2}{2a}. \tag{4}$$

The distance of the following vehicle traveled:

$$S_2 = \frac{V_0^2 - V_2^2}{2a}. \tag{5}$$

For avoiding a rear-end collision between the consecutive vehicles, the following condition shall be met:

$$d > S_2 - S_1 = \frac{V_1^2 - V_2^2}{2a} = \frac{(V_1 - V_2)(V_1 + V_2)}{2a}. \tag{6}$$

A rear-end collision would not happen if the speed of the following vehicle V_2 is smaller than that of the leading one V_1. However, when the speed relationship is reversed, whether the collision would occur will depend on the distance d, the speed difference between the consecutive two vehicles $V_2 - V_1$, the sum of the speed of two vehicles $V_2 + V_1$, and the deceleration rate a. The results in the previous section showed that suitable VSL control can decrease the mean speed, the ASD, and the percentage of relatively small space headway (which means most drivers prefer to keep a large space headway). As a result, under reasonable VSL controls, the speed difference between the two consecutive vehicles V_2-V_1 and the sum of the two vehicles' speeds both become smaller in the right part of (5). Meanwhile, the possibility of space headway d in the left part of (5) is also small. Therefore, it can be inferred that appropriate VSL controls can decrease the risk of rear-end collisions.

3.5. Driver's Compliance under VSL. This section focuses on understanding the driver's compliance under VSL because the driver's compliance determines the efficiency of the VSL. Based on our analysis, it can be observed that a lower VSL value does not guarantee a higher impact on driving behavior. For instance, under the low VSL values, the mean speed, average speed difference, and the percentage of small headways are greater than the same measurements under high VSL values. The reason for this phenomenon is the change in the driver's compliance with the variable speed limit control [8].

Whether or not the driver complies with the VSL is a process of measuring the greater of the benefits and the losses. From the driver's perspective, the benefits that may be gained from complying with the VSL include improved driving safety and no risk of penalties for speeding; the corresponding losses include reduced vehicle speed and increased travel time. On the contrary, if drivers fail to comply with the VSL, the possible benefits are efficiency benefits from higher driving speeds, etc.; the corresponding losses include reduced driving safety and the risk of penalties for speeding. Only when the driver judges that the benefits of the VSL are greater than the losses, they will choose to comply. When the VSL value is too low, that is, when the speed limit is significantly different from the driver's expected speed, the driver may question the rationality of the VSL. Under this situation, the driver's awareness of expecting to reach the destination as soon as possible defeats the awareness of complying with VSL limits to obtain safety benefits. At this time the driver will choose not to comply with the VSL.

Although the driver will choose not to follow the low VSL values, this does not mean that the driving behavior under low VSL values is the same as that without VSL. Our analysis finds that when the VSL values are too low, although no driver will choose a speed slower than the variable speed limit value, the mean speed is still significantly lower than the uncontrolled state. For example, under 60 km/h VSL, the speed of almost all vehicles is higher than 60 km/h, and the

TABLE 4: Percentages of smaller than 100 m headway within different density intervals under different speed limits ("—" means less than 300 vehicles are observed).

Lane	Density (veh/km)	Percentage of less than 100-m headway under different speed limits (%)					
		No VSL (%)	120 km/h (%)	100 km/h	80 km/h	60 km/h	50 km/h
Left lane	5–10	23	21	15%	—	—	—
	10–15	32	28	19%	17%	—	—
	15–20	35	26	15%	18%	14%	—
	20–25	38	27	18%	23%	28%	24%
	25–30	42	33	33%	38%	42%	—
Middle lane	5–10	9	6	—	—	—	—
	10–15	15	11	8%	—	—	—
	15–20	22	17	13%	17%	—	—
	20–25	28	22	19%	23%	35%	21%
	25–30	34	30	29%	—	—	—
Right lane	5–10	4	4	—	—	—	—
	10–15	9	9	8%	9%	—	—
	15–20	16	18	12%	14%	22%	—
	20–25	29	27	—	—	—	—
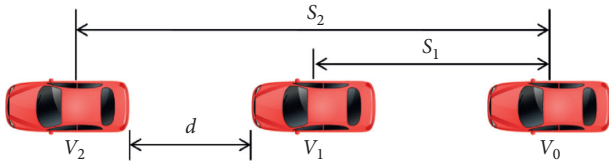
FIGURE 3: Schematic diagram of rear-end collisions.

mean speed is around 85 km/h, while the average speed in the uncontrolled state under the same density condition is around 110 km/h.

In the real world, each driver has an expected speed that considers safety and efficiency under any conditions. Drivers will drive as fast as possible based on ensuring driving safety and complying with traffic laws. The VSL is to change the driver's expected speed by changing the driver's safety expectations and the expectations of complying with traffic laws. Therefore, by popularizing the safety benefits of VSL and strengthening law enforcement for speeding violations, compliance with VSL can be effectively improved, thus improving the overall safety impacts of VSL.

4. Conclusions and Discussion

With the high-resolution lane-by-lane traffic big data collected from a European motorway, this study quantitatively measures the difference in driving behavior under different VSL values and analyzes the safety impacts of VSL controls on aggregate driving behaviors (described by mean speed, ASD, and the percentage of small space headway). Also, the potential improvement of VSL on rear-end collisions and the driver's compliance under VSL was discussed to prove that appropriate VSL controls could have provided additional traffic safety benefits. The main findings of this study are summarized as follows:

(1) Under medium and lower traffic density conditions, different VSL values can always reduce the mean traffic speeds. But the changes in the mean speed followed the first decreases and then increases trend with the reduction in speed limits. The result also indicated that, under low VSL values, drivers are more likely to challenge the rationality of the VSL system.

(2) Under similar traffic conditions, most VSL values reduced the speed differences between consecutive vehicles, thereby reducing the speed discretions in the traffic stream. Again, the first decreases and then increases trend with the reduction in speed limits in ASD showed that low VSL values may increase the ASD.

In general, VSL can reduce the mean speed, the speed difference, and the percentage of small headway. Reasonable VSL control strategies can effectively reduce the risk of rear-end collisions on the motorway. However, the first decreases and then increases trend with the reduction of speed limits illustrated that, under low VSL values, drivers are more likely to challenge the rationality of the VSL system. Therefore, the overly low VSL values cannot provide a higher safety benefit or even caused more traffic accidents. Thus, the implementation of the VSL control system should closely be related to the real-time traffic flow congestion level.

One limitation of this study is that it only considers the impact of VSL on the three aggregated driving behaviors under the same traffic density interval. On the other hand, VSL may have an impact on density under the same traffic demand. Future research is suggested to consider how VSL will impact the traffic density with the same traffic demand. In addition, the data used in this study does not include any lane-changing information. Since the lane-changing rate is another vital index for traffic stability and safety, the impact of VSL on the lane-changing rate is also

encouraged to be analyzed with more detailed data collected.

Data Availability

The basic data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] M. Papageorgiou, *Applications of Automatic Control Concepts to Traffic Flow Modeling and Control*, Springer, Berlin, Germany, 1983.

[2] B. Khondaker and L. Kattan, "Variable speed limit: an overview," *Transportation Letters*, vol. 7, no. 5, pp. 264–278, 2015.

[3] P. Rämä and J. Luoma, "Driver acceptance of weather-controlled road signs and displays," in *Transportation Research Record: Journal of the Transportation Research Board, No.1573*, pp. 72–75, Transportation Research Board of the National Academies, Washington, D.C., USA, 1997.

[4] P. Rämä, "Effects of weather-controlled variable speed limit and warning signs on driver behavior," in *Transportation Research Record: Journal of the Transportation Research Board, No.1689*, pp. 53–59, Transportation Research Board of the National Academies, Washington, D.C., USA, 1999.

[5] C. Lee and M. Abdel-Aty, "Testing effects of warning messages and variable speed limit on driver behavior using driving simulator," in *Transportation Research Record: Journal of the Transportation Research Board, No.2069*, pp. 55–64, Transportation Research Board of the National Academies, Washington, D.C., USA, 2008.

[6] M. Yang, Z. Li, Z. Ke, and M. Li, "A deep reinforcement learning-based ramp metering control framework for improving traffic operation at freeway weaving sections," in *Proceedings of The 98th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA, January 2019.

[7] W. Zhou, M. Yang, M. Lee, and L. Zhang, "A Q-learning based coordinated variable speed limit and hard shoulder running control strategy to reduce travel time at freeway corridor," *Transportation research record: Journal of the Transportation Research Board*, vol. 2674, p. 915, 2020.

[8] X. Qu, L. Li, Z. Yi, P. Mao, and M. Yang, "Traffic flow modeling of freeway variable speed limit control based on the big data of driving behavior," *Journal of Advanced Transportation*, vol. 2020, Article ID 8859494, 11 pages, 2020.

[9] J. R. D. Frejo, I. Papamichail, M. Papageorgiou, and B. D. Schutter, "Macroscopic modeling of variable speed limits on freeways," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 15–33, 2019.

[10] C. Lee, B. Hellinga, and F. Saccomanno, "Assessing safety benefits of variable speed limit," in *Transportation Research Record: Journal of the Transportation Research Board, No.1897*, pp. 183–190, Transportation Research Board of the National Academies, Washington, D.C., USA, 2004.

[11] M. Abdel-Aty, R. Cunningham, V. Gayah, and L. Hsia, "Dynamic variable speed limit strategies for real-time crash risk reduction on freeways," in *Transportation Research Record: Journal of the Transportation Research Board, No.2078*, pp. 108–116, Transportation Research Board of the National Academies, Washington, D.C., USA, 2008.

[12] E. Grumert, X. Ma, and A. Tapani, "Analysis of a cooperative variable speed limit system using microscopic traffic simulation," *Transportation Research Part C: Emerging Technologies*, vol. 52, pp. 173–186, 2015.

[13] X. Qu, W. Wang, W.-f. Wang, and P. Liu, "Real-time rear-end crash potential prediction on freeways," *Journal of Central South University*, vol. 24, no. 11, pp. 2664–2673, 2017.

[14] X. Qu, M. Yang, F. Yang, B. Ran, and L. Li, "An improved single-lane cellular automaton model considering driver's radical feature," *Journal of Advanced Transportation*, vol. 2018, no. 10, Article ID 3791820, 2018.

[15] C. Chen, K. Petty, A. Skabardonis, P. Varaiya, and Z. Jia, "Freeway performance measurement system: mining loop detector data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1748, no. 1, pp. 96–102, 2001.

[16] M. Yang, J. Xie, P. Mao, C. Wang, and Z. Ye, "Application of the ARIMAX model on forecasting freeway traffic flow," in *CICTP 2017: Transportation Reform and Change—Equity, Inclusiveness, Sharing, and Innovation*, pp. 593–602, American Society of Civil Engineers, Reston, VA, USA, 2018.

[17] H. Zackor, "Self-sufficient control of speed on freeways," in *Proceedings of International Symposium on Traffic Control Systems*, pp. 226–249, University of California, Berkeley, CF, USA, December 1979.

[18] S. Smulders, "Control of freeway traffic flow by variable speed signs," *Transportation Research Part B: Methodological*, vol. 24, no. 2, pp. 111–132, 1990.

[19] E. Van den Hoogen and S. Smulders, *Control by Variable Speed Signs: Results of the Dutch Experiment*, IET, in *Proceedings of Seventh International Conference on Road Traffic Monitoring and Control*, pp. pp145–149, IET, London, UK, April 1994.

[20] T. Ha, J. Kang, and J. Park, *The Effects of Automated Speed Enforcement Systems onTraffic-Flow Characteristics and Crashes in Korea*, pp. 28–31, Institute of Transportation Engineers, Washington, DC, USA, 2003.

[21] P. Borrough, *Variable Speed Limit Reduce Crashes Significantly in the UK*, The Urban Transportation Monitor, 1997.

[22] Y. Pilli-Sivola, *State of the Art in Finland Concerning RWIS and Variable Message Signs*, Finnish National Road Administration, Helsinki, Finland, 2000.

[23] M. Papageorgiou, E. Kosmatopoulos, and I. Papamichail, "Effects of variable speed limit on motorway traffic flow," in *Transportation Research Record: Journal of the Transportation Research Board, No. 2047*, pp. 37–48, Transportation Research Board of the National Academies, Washington, D.C., USA, 2008.

[24] F. Soriguera, I. Martínez, M. Sala, and M. Menéndez, "Effects of low speed limits on freeway traffic flow," *Transportation Research Part C: Emerging Technologies*, vol. 77, pp. 257–274, 2017.

[25] H. Wang, W. Wang, X. Chen, J. Chen, and J. Li, "Experimental features and characteristics of speed dispersion in urban freeway traffic," in *Transportation Research Record: Journal of the Transportation Research Board, No.1999*, pp. 150–160, Transportation Research Board of the National Academies, Washington, D.C., USA, 2007.

[26] U. Kohler, "Stability of vehicle platoons," *Transportation and Traffic Theory*, vol. 6, pp. 39–55, 1974.

WILEY | Hindawi

Research Article

Traffic Status Prediction of Arterial Roads Based on the Deep Recurrent Q-Learning

Wei Hao [iD],[1] **Donglei Rong,**[1] **Kefu Yi** [iD],[2] **Qiang Zeng** [iD],[3] **Zhibo Gao** [iD],[4] **Wenguang Wu** [iD],[2] **Chongfeng Wei,**[5] **and Biljana Scepanovic**[6]

[1]*Hunan Key Laboratory of Smart Roadway and Cooperative Vehicle-Infrastructure Systems, Changsha University of Science and Technology, Changsha, Hunan 410205, China*
[2]*School of Automotive and Mechanical Engineering, Changsha University of Science and Technology, Changsha 410114, China*
[3]*School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, China*
[4]*Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai, China*
[5]*Mechanical and Construction Engineering, Northumbria University, Ellison Place, Newcastle upon Tyne NE1 8ST, UK*
[6]*Faculty of Civil Engineering, University of Montenegro, 81000 Podgorica, Montenegro*

Correspondence should be addressed to Kefu Yi; corfyi@csust.edu.cn, Qiang Zeng; zengqiang@scut.edu.cn, Zhibo Gao; gaozhibo@tongji.edu.cn, and Wenguang Wu; wwglq@csust.edu.cn

With the exponential growth of traffic data and the complexity of traffic conditions, in order to effectively store and analyse data to feed back valid information, this paper proposed an urban road traffic status prediction model based on the optimized deep recurrent Q-Learning method. The model is based on the optimized Long Short-Term Memory (LSTM) algorithm to handle the explosive growth of Q-table data, which not only avoids the gradient explosion and disappearance but also has the efficient storage and analysis. The continuous training and memory storage of the training sets are used to improve the system sensitivity, and then, the test sets are predicted based on the accumulated experience pool to obtain high-precision prediction results. The traffic flow data from Wanjiali Road to Shuangtang Road in Changsha City are tested as a case. The research results show that the prediction of the traffic delay index is within a reasonable interval, and it is significantly better than traditional prediction methods such as the LSTM, K-Nearest Neighbor (KNN), Support Vector Machines (SVM), exponential smoothing method, and Back Propagation (BP) neural network, which shows that the model proposed in this paper has the feasibility of application.

1. Introduction

With the development of urbanization, there is a prominent contradiction between the transportation infrastructure and the vehicle population, and the problem of traffic congestion has become more serious, which inevitably leads to the increasing of travel time, intensified environmental pollution, and economic loss [1]. Prevention is the first way to control traffic congestion. According to the existing traffic states, the changing trend in a short time is predicted, and then, the information platform is used to issue an early warning to divert the traffic to avoid or ease congestion

[2–4]. Therefore, how to establish a long-term model for timely warning of traffic congestion is the research focus of urban intelligent transportation system optimization [5–7].

A variety of methods, including the time series, machine learning, and artificial neural networks, have been proposed for traffic congestion prediction. Since the time-series characteristics of traffic flow data were discovered [8], some scholars used autoregressive differential moving average models [9] to predict the traffic flow on expressway [10, 11]. Because the temporal distribution of traffic flow data is interrelated, some scholars used nonparametric regression methods to build macrotraffic models and found that the

prediction result is better than time-series algorithms [12–14]. However, these methods based on statistics and traffic models require a large amount of historical data and construct many assumptions, so they are difficult to apply to nonlinear traffic flow [15–18].

In recent years, machine learning algorithms, such as the back propagation neural network [19, 20], have gradually been used in traffic prediction with the advantage to handle the nonlinearity problems. Because of the long training time of the back propagation neural network and the tendency to fall into the local optimum, some scholars also used the Support Vector Machine (SVM) [21–23] and K-Nearest Neighbor (KNN) [24–26] to predict the traffic status. Moreover, some scholars found that the time series of short-term traffic flow has chaotic characteristics. To deal with the abovementioned issues, many methods, such as combined vector machine-based [27] and phase space reconstruction-based [28], have been proposed to achieve better results. However, most of these machine learning-based methods lack robustness to catch the huge data, resulting in the model generally lacking long-term effectiveness and scalability [29–31].

Facing on the lots of traffic flow data, scholars have gradually turned to use the deep learning method, a learning algorithm that can simulate the multilayered perceptual structure of the human brain to recognize the data patterns. At present, breakthroughs have been made in many fields such as computer vision, speech recognition, and natural language processing. Deep learning has gradually been adopted by Stanford University, Google, Baidu Research Institute, and other authoritative organizations with the strategic direction for the development of data mining and artificial intelligence [32, 33]. Kuremoto et al. [34] combined the restricted Boltzmann machine with the time-series laws to obtain a prediction model, which fits the sample data with the minimum model energy. Lv et al. [35] proposed a deep learning model to predict traffic flow based on an automatic coding network using compression coding in the input data. Zhao et al. [36] proposed a traffic congestion prediction model based on the improved SVM, which can learn the characteristics of traffic flow parameters through the deep structure by digitizing different environmental and human factors. The abovementioned methods speed up data processing by applying the deep learning models but do not take into account the dimensional disaster caused by the high-dimensional states of traffic flow parameters. To address the abovementioned problems, some scholars used data compression technology based on the LSTM, Principal Component Analysis (PCA) [37], CUR matrix decomposition algorithm [38], and Discrete Cosine Transform (DCT) method [39] to perform data dimension reduction.

Q-Learning can efficiently store and extract data to provide support for traffic prediction. The LSTM network reduces the frequency of gradient explosion and disappearance, so it is suitable for capturing the spatiotemporal evolution of traffic state parameters [40–43]. In this paper, considering the time sequence of traffic flow parameters and the continuity of traffic congestion effects, the recurrent neural network model is used to train the extracted features

and to obtain low-dimensional vectors of historical information, and then, the resulting vectors are stitched to achieve classification training. Finally, an urban road traffic state prediction model based on the optimized deep recurrent Q-Learning method is established. The model proposed in this paper has the following contributions:

(1) The model effectively solves the problem of gradient explosion and gradient disappearance in the prediction process of LSTM

(2) The model effectively extracts the associated features of the traffic data, so it has better prediction efficiency and accuracy

(3) The model will provide a feasible prediction method for the construction of an intelligent transportation system due to its efficiency and feasibility

The rest of this paper is organized as follows. Section 2 points out the problems to be solved and the corresponding methods in this paper. Then, Sections 3 and 4 lead to the principles and steps of the Q-Learning and the LSTM. After that, the deep recurrent Q-Learning network model is constructed in Section 5. Besides, the example analysis in Section 6 proves the stability and feasibility of the method. Finally, Section 7 concludes the paper.

2. Specific Problems and Solutions

2.1. Specific Problems. The problems with urban traffic data are high repeatability, high loss rate, and poor correlation. The existing prediction methods mainly discuss the results of independent analysis and whether they meet the needs of further verification. Therefore, the following problems exist in data preprocessing and optimization prediction.

Regarding the problem of data relevance: the relationship between the states at the previous moment and the next moment lacks effective connection. Therefore, the information at different states is disconnected, and the timeliness of the data cannot be fully exerted. As a result, the prediction results are not sufficiently correlated with the data at the previous moment and lack of persuasiveness.

Speaking of the problem of data storage: based on the existing analysis methods, the storage capacity of the database will quickly reach the threshold, which is not conducive to long-term and durable prediction. Besides, repeated analysis steps will increase the feedback delay and cannot fulfill the requirements of low-latency traffic prediction.

Concerning the problem of comprehensive data analysis: the existing analysis focuses on fixed types of data, and the traffic environment is an integrated system. Therefore, even if the prediction results are accurate, they cannot reflect the objective situation.

2.2. Solutions. For the abovementioned three research problems, this paper will propose the corresponding solutions:

For the problem of data relevance: based on the optimized LSTM model, the effective correlation and

information accumulation of different data types are strengthened, and the correlation degree of data at different moments is strengthened.

For the problem of data storage: the Q-Learning functionalizes the data information, and each data cell can be realized by the expression of functions. This method not only reduces the pressure of data storage but also improves analysis efficiency and accuracy.

For the problem of comprehensive data analysis: the traffic conditions are affected by multiple factors. Therefore, when selecting the characteristic data types, in addition to the basic parameters of traffic flow, climate and temperature are also considered. That is to establish a multidimensional data analysis system, making the prediction results more accurate and objective.

3. Q-Learning Principle and Application Steps

The steps of the Q-Learning are listed as follows: the state of the agent in the environment E is S, and the actions taken by the agent constitute the action space A. It takes different actions to transfer between states, and the reward function obtained is R. To achieve the optimal strategy, the Q-Learning estimates the value of each action choice in each state. The Q-Learning uses $Q(S, A)$ to represent the value function of state-to-action and continuously updates the value of $Q(S, A)$ according to the state transition. Finally, the Q-Learning obtains the optimal strategy based on $Q(S, A)$.

The value function $Q(S, A)$ of the traffic state is updated as follows: assuming the state of the agent at time t is s, the action is a_t. Then, the state transitions to time $t + 1$, the state is s_{t+1}, and the reward is r_t. Finally, the agent updates the value of $Q(s, a)$ according to all records (s_t, a_t, t_t, s_{t+1}) to find the optimal strategy. The corresponding update function is shown in the following equation:

$$Q(s_t, a_t) \longleftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \lambda \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right],$$
(1)

where $Q(s_t, a_t)$ is the current Q-table, α is the learning rate, r_{t+1} is the benefit at the next moment, λ is the greed coefficient, and $\max_a Q(s_{t+1}, a_t)$ is the best benefit in memory.

The deep Q-Learning network combines deep learning and Q-Learning. The network uses the perceptive ability of the deep learning to transform the state to high dimensions and uses the decision-making ability of Q-Learning to map the high-dimensional state representation to the low-dimensional action space [44, 45]. In the Q-Learning algorithm, the table is used to store the value of $Q(s, a)$. In the deep Q-Learning, the state dimension of the agent is high, and the table obviously cannot meet the demand. This problem is solved by using $f(s, a)$ to approximate $Q(s, a)$ [46, 47]. Therefore, based on the corresponding value function neural network model, approximate values can be obtained, thereby reducing the storage pressure of the Q-table and providing ideas and methods for Q-Learning to be applied to traffic state prediction. Finally, the network obtains the action value of congestion and dissipation according to the accumulated experience pool. Figure 1

shows a schematic diagram of the principle of approximating the value of "state-action" through the neural network.

The network helps solve the problems of processing huge data volumes. Due to the strong time series of traffic data, the application of this network will make the analysis results more reliable. Further demonstrations and experiments will be discussed in the following sections.

4. Recurrent Neural Network LSTM Algorithm

4.1. Overview of the Recurrent Neural Network. The recurrent neural network is one of the optimized variants of deep neural networks. It is characterized by the output of the neurons at a certain moment as part of the input of the next moment, and the neural network has the function of memorizing the information of the previous moment which can realize the persistence of the information. As shown in Figure 2, the neural network reads the input x_t of the current time t and obtains the output h_t. At the same time, the information status is returned to the neural network as one of the inputs at the next time point. In order to show the execution action more intuitively, we express it by

$$h_t = f(h_{t-1}, x_t).$$
(2)

The output h_t at each moment is related to the input h_{t-1} at the previous moment. The recurrent neural network is the most natural structure for processing sequence data which is exactly what we need to handle historical data and real-time data in this paper.

x_t: the input at time t, $X = [x_1, \ldots, x_{t-1}, x_t, x_{t+1}, \ldots, x_T]$ is the input sequence.

s_t: the state of the hidden layer at time t, also known as the memory unit of the recurrent neural network.

h_t: the output at time t, $H = [h_1, \ldots, h_{t-1}, h_t, h_{t+1}, \ldots, h_T]$ is the output sequence.

U: the weight parameter matrix of input sequence information X to hidden layer state S.

W: the weight parameter matrix between the hidden layer states S.

V: the weight parameter matrix of hidden layer state S to output sequence information H.

4.2. Recurrent Neural Network LSTM. If the dependency interval between sequences is long, the gradient disappearance of traffic data will happen in ordinary RNN which is difficult to retain the information at earlier times. The LSTM network remembers long-term historical information through the design of the network structure where the output of the network at time $t + 1$ is applied to itself at time t to avoid the gradient disappearance. Its network expands along the time axis. The schematic diagram and the detailed diagram of the three-layer gate are shown in Figures 3 and 4.
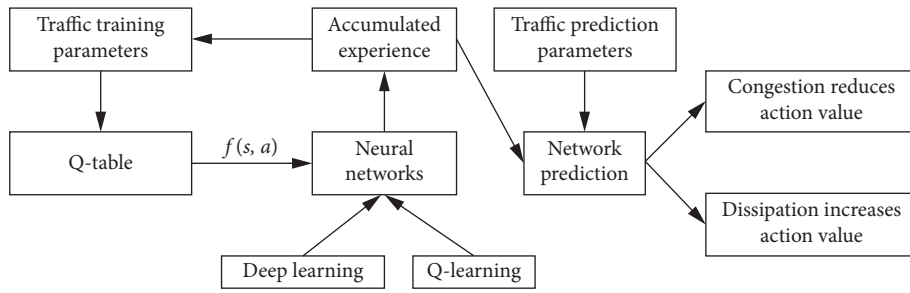
Figure 1: Schematic diagram of the neural network approximate establishment of "state-action" value.
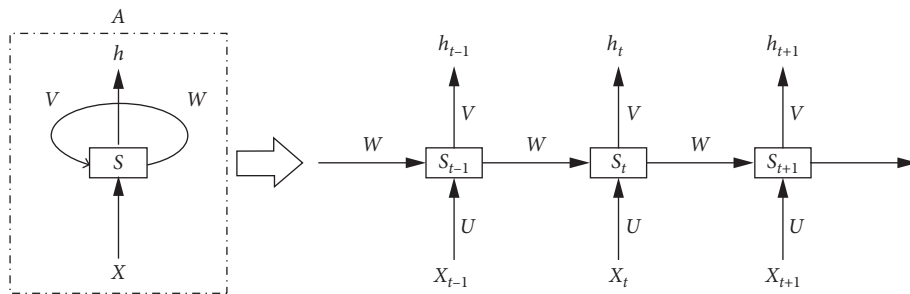
Figure 2: Chain structure of the recurrent neural network.
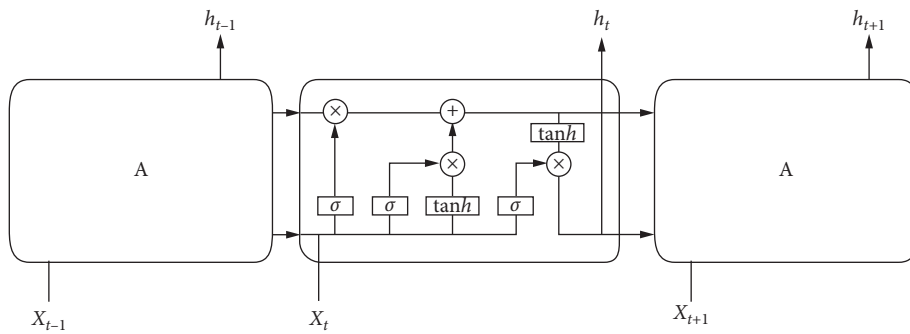
Figure 3: The schematic diagram of LSTM expansion by the time axis.
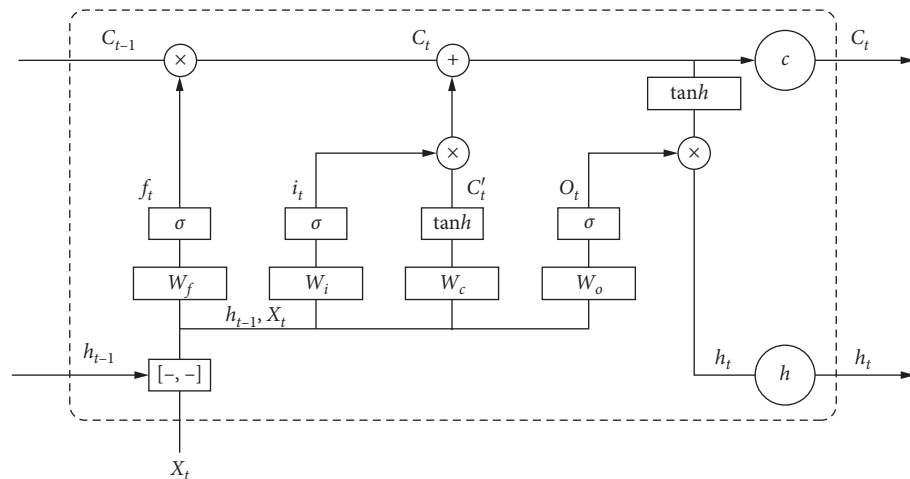
Figure 4: Detailed diagram of three-gate levels.

It can be seen from Figure 3 that the LSTM defines the key concept of cell state with the horizontal line. There is less information interaction in the cell with the purpose of memorizing long-term information achieved through cell transmission. For Figure 4, it is made by three-gate layers with the first one is the forget gate. This gate is determined based on the input of the current moment and the output of the previous moment and, then, passes through a *Sigmoid* layer to obtain the results. It determines how much of the cell states from the previous moment is retained to the current moment. The expression function of the forget gate is shown in the following equation:

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big), \tag{3}$$

where f_t represents the output of the forget gate and σ represents the *Sigmoid* function. W_f and b_f represent the weight matrix and the bias term, respectively. $[h_{t-1}, x_t]$ represents the connection of two vectors into a longer vector.

The second one is the input gate, which determines how much of the network input is saved to the cell state at the current moment. The expression functions of the input gate are shown in the following equations:

$$i_t = \sigma(W_t \cdot [h_{t-1}, x_t] + b_t), \tag{4}$$

$$\widetilde{c}_t = \tan h(W_c \cdot [h_{t-1}, x_t] + b_c), \tag{5}$$

where c_t is calculated by multiplying the last element state c_{t-1} by the element forget gate f_t, then multiplying the element state c_t by element by the input gate i_t, and finally, adding the two products.

The cell information can be updated based on the results of forget gate and output gate. It is listed as follows:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \widetilde{c}_t. \tag{6}$$

The last one is the output gate, which controls how many cell states output to the current output value of the LSTM. From Figure 4, the output gate is composed of two parts, one is the state of the cells processed by tan*h*, and the other is the input information processed by *Sigmoid*. The functions of the output gate are listed in the following equations:

$$\sigma_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \tag{7}$$

$$h_t = o_t \cdot \tan h(c_t), \tag{8}$$

where o_t represents the output of the output gate. W_o and b_o represent the weights and offsets, respectively.

5. Deep Recurrent Q-Learning Network

5.1. State Space. If the amount of acquired data is not large, the Q-Learning can perform data storage and processing efficiently. If the data is large, Q-Learning cannot traverse all states, and there is no such large space to install the Q-value table in memory. Therefore, this paper uses the LSTM model to generalize the states and uses the recurrent neural network to learn the state function. Through continuous deep reinforcement learning, the model obtains features to describe the current state, while accumulating experience pool. In constructing the state space, it is divided into two steps: state discretization and value evaluation.

> Step 1: what the neural network wants to output is the training value under each state, which represents the measure of the pros and cons of developing from this state. The characteristics of the current state S are speed v, delay time d, travel time m, temperature t, and precipitation probability p. If the characteristics of the next state S' are speed $v\prime$, delay time d', travel time m', temperature t', and precipitation probability p', then the corresponding selection behavior of reward accumulation is [speed v, delay time d, travel time m, temperature t, precipitation probability p] minus [speed $v\prime$, delay time d', travel time m', temperature t', precipitation probability p']. The resulting values for each position are positive 1 and negative 0 to discretize the behavior (the range is $0 \sim 2^5$, and the selected behavior discretization vector $[0, 1, 0, \ldots]$ is transformed into an integer which represents the dimension of the output vector). Based on the results of the abovementioned rewards, the traffic information with better benefits is accumulated to form an experienced pool with high benefit values, which makes the prediction results more accurate.

> Step 2: due to the influence of the traffic states before and after the training, it is necessary to determine whether the action can get excellent feedback before execution. The action a is performed according to the strategy p, and the cumulative return is calculated, after the strategy is executed. The state value function expression is listed as follows:

$$V_p(s) = \sum_{s'} p_{(s)}(s, s')\big(R_{p(s)}(s, s') + \gamma V_p(s')\big), \tag{9}$$

where $V_p(s)$ represents the degree of return according to the strategy p under state s. $p(s, s')$ represents the probability of state transition. $R(s, s')$ represents the reward obtained from $s \longrightarrow s'$, and γ is a function coefficient.

5.2. Reward Actions. The reward after training in the previous state s is represented by the difference in delay time. The neural network function uses s as input. $Q(a, n\_features)$ is the storage table, and $n\_features$ presents the number of input neurons. Therefore, the output vector dimension is $2^{n\_features} = 32$. The memory storage pool structure after the reward is $[n\_features, a, r, n\_features]$.

During the process of predicting the future situation, inputting the current state and outputting the Q-value are studied under various possibilities with the largest one being selected. As the reward level continued to deepen, the target results gradually approach the actual situation in which the Q-value here refers to the traffic delay index.

5.3. Training Method. Based on the construction of the abovementioned state space and reward actions, we will train the datasets from Wanjiali Road to Shuangtang Road on the elevated Wanjiali Road in Changsha. The main steps of training methods are listed as follows:

> Step 1: the preprocessing of traffic data and weather data (culling abnormal data, Lagrange interpolation, and normalization).
>
> Step 2: the selection of training sets and test sets (the time interval of training sets is from 0:00 on May 17, 2019, to 24:00 on May 24, 2019. The time interval of test sets is from 0:00 on May 25, 2019, to 12:00 on May 25, 2019).
>
> Step 3: determining the input and output of the variables and the number of network layers (the input variables are speed, delay time, travel time, temperature, and precipitation probability. The output variable is the delay index, the number of hidden neurons in the interval [4, 13], and 3 layers of network layers).
>
> Step 4: determining the initial weights, thresholds, learning rate, activation function, and training function (the interval of initial weight and threshold is [0, 1]. The learning rate is 0.01, the activation function uses the *Sigmoid* function, and the training function uses *Adam*).
>
> Step 5: training the neural network model and stopping the network training when the feedback reaches the optimal state of the *Q*-value table. If it is not satisfied, modification and adjustment of the parameter values are required (learning rate and training function).
>
> Step 6: adjusting the parameter to achieve the best prediction results which could be obtained from the prediction and input test set data.
>
> Step 7: analyzing the prediction results to get the final experimental results.

In this paper, the LSTM forgetting, input, and output threshold activation functions are all *Sigmoid* functions. The return interval [0, 1] is consistent with human thinking. The pseudocode to build a deep recurrent Q-Learning network is shown in Algorithm 1.

6. Case Analysis

6.1. Data Description. This paper selected a part of the arterial road in Changsha, starting from Wanjiali Road to Shuangtang Road from north to south, as the research case. A crawler script written in Python 3.7 was used to capture the real-time traffic information from the big data platform of Gaode Map. The data were collected from 0:00 on May 17, 2019, to 12:00 on May 25, 2019, with a 5-min sampling interval. The collected data types include actual time, speed, delay time, travel time, temperature, probability of precipitation, and delay index. The data set sample is shown in Table 1.

The data of this case is divided into training sets and test sets after preprocessing. The time interval of the training sets is from 0:00 on May 17, 2019, to 24:00 on May 24, 2019, and the time interval of the test sets is from 0:00 to 12:00 on May 15, 2019.

6.2. Data Preprocessing. Data preprocessing includes three steps: culling abnormal data, Lagrange interpolation, and normalization. The detailed information is shown in Figure 5.

The first step is to cull abnormal data. The abnormal data mentioned in this step refers to the data that deviates significantly from the normal interval. By deleting such kind of data, the experimental data are more realistic and the analysis results are more reasonable. Some samples of abnormal data are shown in Table 2.

The second step is the Lagrange interpolation. Lagrange interpolation is used to fill in some missing data based on the neighboring traffic datasets to improve the value of the data. This step is used to achieve data integrity and rationality.

The data filling function for this step is listed as follows:

$$L(x) = \sum_{i=0}^{n} y_i \prod_{j=0, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}, \qquad (10)$$

where y_i is the polynomial of degree $i-1$ and x_i is the parameter corresponding to the point i.

The third step is data normalization. The purpose of this step is to control the magnitude of the data within a small fluctuation range, reduce the impact between the magnitudes of the horizontal data, and improve prediction accuracy. The function is listed as follows:

$$x' = \frac{x - \min}{\max - \min}, \qquad (11)$$

where max is the maximum value of the sample data and min is the minimum value of the sample data.

The preprocessed data are transformed into a list to form the matrix and, finally, transformed into a three-dimensional space. The three-dimensional space serves as the input of the LSTM unit to form the basic unit of the hidden layer. Every 15 rows of valid data are used as a training set and are continuously trained in 100 times. The test sets are predicted based on the training memory to obtain the prediction results.

6.3. Prediction Module Construction. According to data analysis, visualization, and platform requirements, this paper introduces NumPy, Pandas, and Matplotlib as analysis tools. Tensorflow is used as an open-source database for deep learning to build a basic library. According to the needs of the model, a variety of modules is constructed including the traffic environment module, deep reinforcement learning module, memory module, behaviour selection module, neural network module, training main program module, loss curve module, and visualization module.

The first step is to initialize the traffic data, network environment, and training parameters to build a neural network for prediction. The second step is to input training sets and test sets to the input layer. Multidimensional data

(1) Initial network structure, the parameter is q. Initial target network, parameter $q' = q$.
(2) Initial trials greedy parameters epsilon, learning rate, reward, attenuation coefficient gamma, number of iterations episodes. Each episode iteration round number T, training batch size, and neural network parameter rotation cycle transfer_cycle.
(3) for an episode in Episodes do
(4)   Initial traffic state $s_t = s_0$
(5)   For t from 0 to T:
(6)   Selection behavior. (Output an integer with a range of 0 to $2^{n\_features-1}$): Select $a_t = \text{argmax}_a Q(s, a, \theta)$ with a probability of 1-epsilon, and randomly select the behavior a_t with a probability of epsilon.
(7)   After the behavior is determined, find all states s_{all} in the data table that match this behavior, and then randomly select one from s_{all} as s_{t+1} (If no match is found in s_{all}, the behavior is redetermined).
(8)   Put experience (s_t, a_t, r_t, s_{t+1}) into the memory pool.
(9)   Take out batch size data randomly and calculate q_eval and q_next respectively.
(10)   Construct: $y = r_1 + g\max_{t+1} Q(s_{t+1}, a \mid q) \longrightarrow q\_target$
(11)   According to q_eval and q_target, back propagation to improve the network q.
(12)   If the number of iterations is an integer multiple of transfer_cycle, then updates $q' = q$.
(13)   Current state $= s_{t+1}$.
(14)   When the maximum iteration number T of a single round game is reached, the training of this round is stopped, and the traffic state is returned to the initial trial.
(15)   end for
(16) end for

ALGORITHM 1: Algorithm pseudocode.

TABLE 1: Example of traffic state parameters returned from the Gaode big data platform.

Time frame	Speed (km/h)	Travel time (min)	Delay time (min)	Temperature (℃)	Precipitation	Delay index
2019/5/17 3:23	61	16.2	2.1	18	0.2	1.5

Note. The road delay index is used as an evaluation index for the urban congestion degree, which refers to the ratio of actual travel time to the free-flow travel time for the city residents.
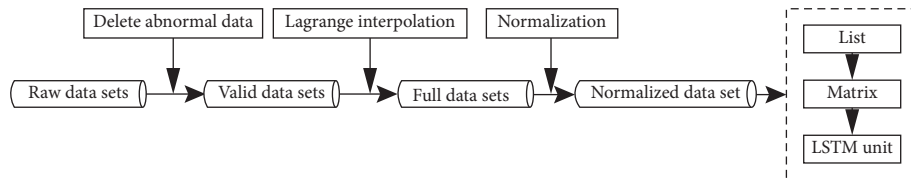
FIGURE 5: Data preprocessing flowchart.

TABLE 2: Example of abnormal data.

Time frame	Speed (km/h)	Travel time (min)	Delay time (min)	Temperature (℃)	Precipitation	Delay index
2019/5/18 17:44	55	14.3	2.1	**50**	0.2	1.5
2019/5/18 20:09	**115**	10.2	3	20	0.15	1

introduction is performed in the hidden layer, and data prediction is performed based on the experience pool in the output layer. Meanwhile, the structural dimensions of the input and output are displayed at each stage. The detailed flowchart is shown in Figure 6.

6.4. Parameter Impact Analysis.
For the traffic prediction, the most critical indicators are prediction efficiency and accuracy. Therefore, the parameter impact analysis, the optimization index analysis, and the accuracy analysis are performed in the following section. The parameter impact analysis and optimization index analysis are used for the

evaluation of prediction efficiency, and the accuracy analysis is utilized for the evaluation of prediction accuracy.

During the neural network of traffic prediction, the key parameters affecting the efficiency of the experiment are studied including learning rate, reward decay, greedy, memory size, replacement interval, and batch size. The batch size is the fixed parameter, and the remaining items are variable parameters. Group 6 is the initial parameter group for comparison and analysis with other groups. Therefore, the parameter groups are divided into weakened state parameter groups (group 1–5) and strengthened state parameter groups (group 7–11).
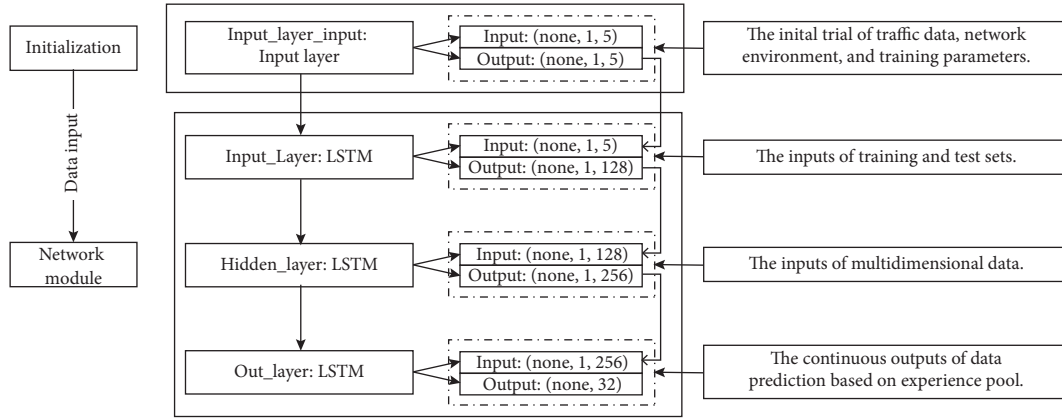
FIGURE 6: Model execution flowchart.

Each group only weakens or strengthens one parameter for comparison with group 6. To improve the discrimination of the experimental results, the selection of parameter values has an obvious gradient with the specific values which is shown in Table 3.

In order to reflect the differences in the experimental results of each group, this paper selected the indicators with obvious discrimination in the experimental results for analysis. They are the highest loss index, the lowest loss index, the maximum volatility, training time, prediction time, and total time. The detailed index distribution of each group is shown in Table 4.

The highest loss index and maximum volatility of group 4 have serious deviations from other groups and have exceeded the normal fluctuation range. It shows that the experimental efficiency analysis in group 4 has no research value. Therefore, group 4 is eliminated before performing a comparative analysis.

The qualitative analysis is performed first. Since the highest loss index, the lowest loss index, the largest volatility rate, and the total time are important parameters of experimental efficiency, visualization is performed, as shown in Figure 7.

From this, the following conclusions are reached:

(1) The indicators of group 6 are all at the highest level, so the gradient weakening or gradient strengthening of the parameters can optimize the experimental results, but pay attention to the combination of extreme parameters. The prediction result at this time does not have actual value, such as group 4.

(2) With the gradient adjustment of parameters, all indicators can fluctuate within a relatively small range without a sharp rise or a sharp decline. Therefore, the stability of the model proposed in this paper is confirmed.

(3) The weakened and strengthened states of memory size and replacement interval have only slight fluctuations compared to the initial state, indicating that these two types of parameters have a little effect on the experimental efficiency.

The quantitative analysis is performed next. According to the optimization degree of each parameter, the effect of improving experiment efficiency is determined. The parameter optimization degree function is shown in the following equation:

$$O = \sum_{i=1}^{4} \frac{x_0 - x_i}{x_0} \times w_i, \qquad (12)$$

where O represents the optimization degree of each group, x_0 represents the parameter value of the initial group, x_i represents the parameter value of the variable group, and w_i represents the optimization weight of the corresponding parameter.

From the perspective of forecasting efficiency, the maximum volatility and total time are representative, followed by the highest loss index and the lowest loss index. Therefore, the initial weight distribution of each parameter is shown in Table 5.

Based on equation (12) and weight distribution, this paper performed optimization calculations for the weakened group compared to the initial group and the strengthened group compared to the initial group.

The following conclusions are drawn based on the quantitative results in Table 6.

(1) The optimization effect of the groups is adjusted by the memory size and replacement interval is weaker than other groups, which further confirms the conclusion (3) in the qualitative analysis.

(2) The optimization effects of all parameters on the experimental efficiency are quite different which indicated that there is parameter emphasis.

(3) The weakened or strengthened groups improve the experimental efficiency, indicating that group 6 is already at or near the worst parameter combination. From this, the lowest limit of the parameter combination can be determined.

This section analyses the effect of five parameters on the experimental efficiency. The results show that memory size and replacement interval have a small effect on experimental

TABLE 3: Gradient parameter value table.

| States | Group | Variable parameters | | | | | Fixed parameter |
		Learning rate	Reward decay	Greedy	Replacement	Memory size	Batch size
Weakened states	1	0.03	0.6	0.6	200	**300**	32
	2	0.03	0.6	0.6	**100**	400	32
	3	0.03	0.6	**0.3**	200	400	32
	4	0.03	**0.3**	0.6	200	400	32
	5	**0.01**	0.6	0.6	200	400	32
Initial state	6	0.03	0.6	0.6	200	400	32
Strengthened states	7	**0.05**	0.6	0.6	200	400	32
	8	0.03	**0.9**	0.6	200	400	32
	9	0.03	0.6	**0.9**	200	400	32
	10	0.03	0.6	0.6	**300**	400	32
	11	0.03	0.6	0.6	200	**500**	32

TABLE 4: Experimental result table.

Group	Highest loss index	Lowest loss index	Maximum volatility	Training time (s)	Prediction time (s)	Total time (s)
1	0.240	0.082	0.0400	1080	100	1180
2	0.230	0.066	0.0385	1088	102	1190
3	0.195	0.070	0.0375	950	92	1042
4	**0.970**	**0.096**	**0.4350**	**1086**	**101**	**1187**
5	0.154	0.070	0.0435	1096	102	1198
6	0.250	0.080	0.0600	1139	105	1244
7	0.186	0.062	0.0510	1000	93	1093
8	0.195	0.080	0.0400	1033	96	1129
9	0.157	0.065	0.0440	1046	99	1145
10	0.238	0.070	0.0400	1013	94	1107
11	0.235	0.081	0.0440	1024	96	1120
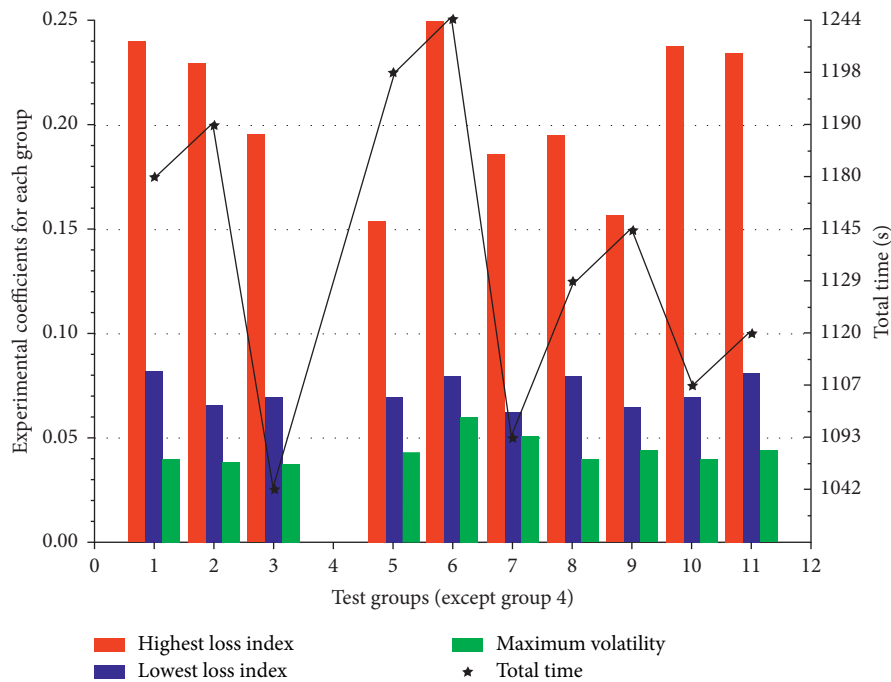
FIGURE 7: Comparison chart of highest loss index, lowest loss index, and maximum volatility.

TABLE 5: Parameter weight distribution table.

Parameter	Highest loss index	Lowest loss index	Maximum volatility	Total time
Weights	20%	20%	30%	30%

TABLE 6: The optimization situation table (except group 4).

The weakened group compared to group 6				The strengthened group compared to group 6				
1	2	**3**	5	7	8	9	10	11
0.1184	0.1215	**0.2302**	0.1954	0.1776	0.1717	0.2158	0.1276	0.1194

efficiency, while learning rate, reward decay, and greedy have a significant degree of tendency to experimental efficiency. Therefore, the three parameters are analysed for efficiency in the next section.

6.5. Optimization Index Analysis. In this section, we performed optimization index analysis on learning rate, reward decay, and greedy. Based on the abovementioned analysis, the replacement interval is set to 300 and the memory size is set to 500.

This part used the orthogonal experiments of three factors and three levels for evaluation. The three factors are learning rate, reward delay, and greedy, record as $A/B/C$. The corresponding levels are $A_1/A_2/A_3$, $B_1/B_2/B_3$, and $C_1/C_2/C_3$, respectively. Fix A and B at the levels of A_1 and B_1, and match three levels of C with $A_1B_1C_1$, $A_1B_1C_2$, and $A_1B_1C_3$. If $A_1B_1C_3$ is optimal, fix the C_3 level. Then, let A_1 and C_3 be fixed, and match two levels of B with $A_1B_2C_3$ and $A_1B_3C_3$, respectively. After the tests, if $A_1B_2C_3$ is optimal, fix two levels of B_2 and C_3, and try two tests $A_2B_2C_3$ and $A_3B_2C_3$. If $A_3B_2C_3$ is optimal, it is the best level combination.

When the loss curve is more stable and the optimal loss coefficient is lower, the corresponding training is better. The stability of the loss curve is reflected by the amplitude of the curve fluctuation, and the curve formed by the ratio of the loss difference to the time difference can be visually seen. The optimal loss coefficient is obtained directly from the experimental results.

The parameter combinations and corresponding results for the first test are shown in Table 7 and Figure 8.

The first test adjusts the values of greedy. Greedy's weakening adjustment significantly improves the stability of prediction and the optimal loss parameters. Therefore, reasonable adjustment of greedy helps optimize the fluctuations of the network and controls the training within a reasonable range. Based on the abovementioned experimental results and analysis, it is found that the optimal combination in the first test is group c.

The parameter combinations and corresponding results for the second test are shown in Table 8 and Figure 9.

The second test evaluates the changes in the reward delay. The effect of the reward delay on the stability of the system is more significant than that of greedy, indicating that the system is more sensitive to the reward delay. Therefore, the parameter tuning of this indicator can be perfectly combined with greedy to achieve the optimal stability of the

TABLE 7: Parameter combinations for the first test.

Group	Learning rate	Reward delay	Greedy	Optimal loss coefficient
a	0.01	0.9	0.9	0.075
b	0.01	0.9	0.6	0.074
c	**0.01**	**0.9**	**0.3**	**0.067**

system with the optimal combination in the second test in group c.

Parameter combinations and corresponding results for the third test are shown in Table 9 and Figure 10.

The object of the third test tuning is the learning rate, which not only puts forward higher requirements for system stability but also realizes the optimization progress of the loss index. Therefore, the learning rate is one of the three factors which have the greatest impact on the system. Meanwhile, the optimal loss index does not decrease with the increase in the learning rate, which indicates that the system optimization has a threshold and is not negatively correlated. The optimal combination for the third test is group f.

Therefore, the best combination obtained in all the tests is group f. The following conclusions are drawn based on the analysis of experimental results:

(1) The learning rate, greedy, and reward delay affect the stability of the system, among which the reward delay has a greater impact. The learning rate is the only effective parameter to improve the optimal loss index.

(2) The values of the three parameters have corresponding valid intervals. When the interval is exceeded, the prediction process fluctuates sharply and affected experimental efficiency.

(3) No extreme fluctuation occurs during the training and prediction process. Even if there are fluctuations, they are always within a reasonable range.

6.6. Accuracy Analysis. The accuracy analysis is divided into two stages: the comparison of the predicted delay index and the actual delay index and the accuracy analysis of the traditional methods and the method proposed in this paper.

The first stage: the comparison between the predicted delay index and the actual delay index. Taking group f as the
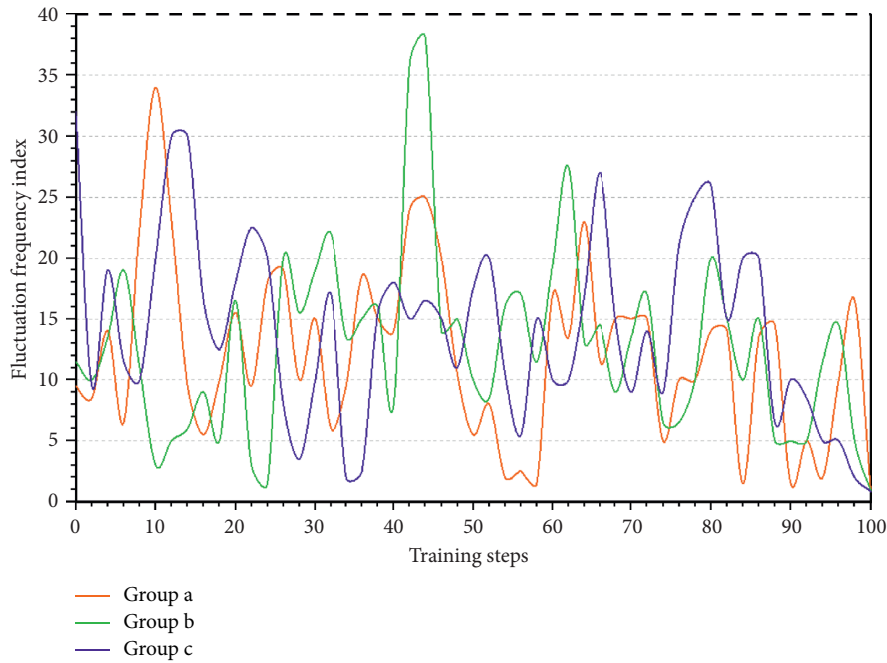
FIGURE 8: Fluctuation frequency graph of groups a, b, and c.

TABLE 8: Parameter combinations for the second test.

Group	Learning rate	Reward delay	Greedy	Optimal loss coefficient
c	**0.01**	**0.9**	**0.3**	**0.067**
d	0.01	0.6	0.3	0.069
e	0.01	0.3	0.3	0.070
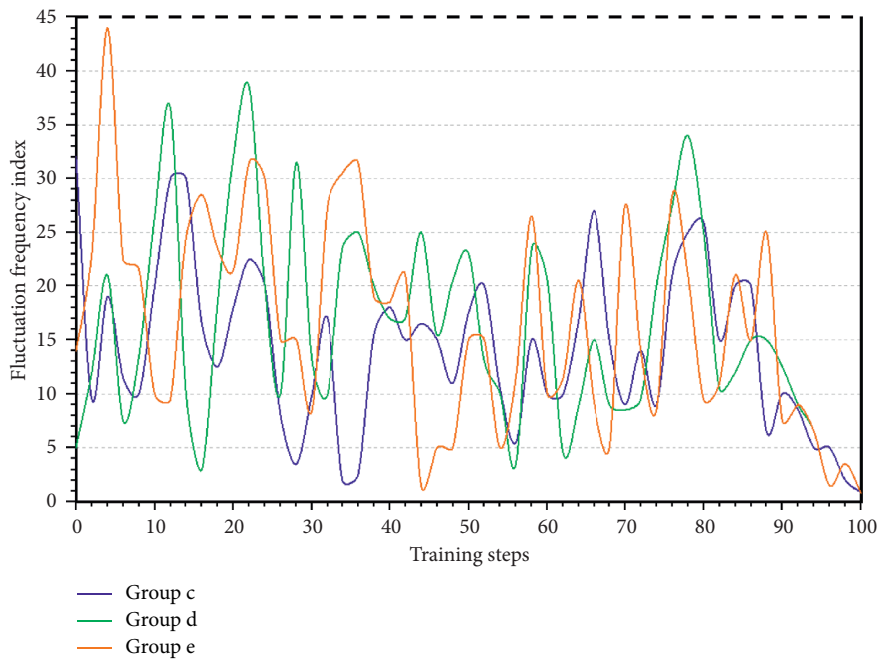
FIGURE 9: Fluctuation frequency graph of groups c, d, and e.

TABLE 9: Parameter combinations for the third.

Group	Learning rate	Reward delay	Greedy	Optimal loss coefficient
c	0.01	0.9	0.3	0.067
f	**0.03**	**0.9**	**0.3**	**0.060**
g	0.05	0.9	0.3	0.061
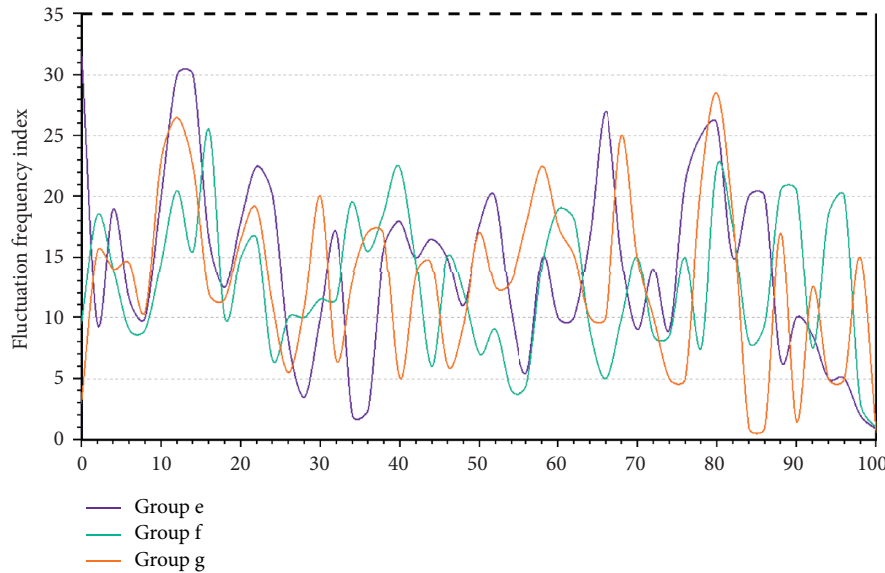
FIGURE 10: Fluctuation frequency graph of groups c, f, and g.

standard group, first draw a comparison chart based on the prediction and actual delay index, and second, calculate the prediction accuracy of groups 1–11 and A–G. This enables a preliminary accuracy evaluation.

It is known from Figure 11 that the degree of agreement between group f and the actual delay index is extremely high, and the predicted efficiency is better in the first half than in the second half. It is known from Tables 10 and 11 that group f still has the highest prediction accuracy. Therefore, the following conclusions are obtained in the first stage:

(1) The neural network under group f is the best choice in terms of experimental efficiency and accuracy, which provides a strong guarantee for short-term traffic prediction.

(2) There is a loss of weakness in the prediction process. If the neural network is used for long-term prediction, the network needs to be further optimized for design.

(3) There is no absolute correlation between experimental efficiency and accuracy, so it is necessary to analyse by yourself. For example, there is a serious deviation in the experimental efficiency of group 4, and its accuracy is kept within a reasonable range.

The second stage: the accuracy analysis of the traditional method and the method mentioned in this paper. In order to further verify the superiority of the proposed method, the
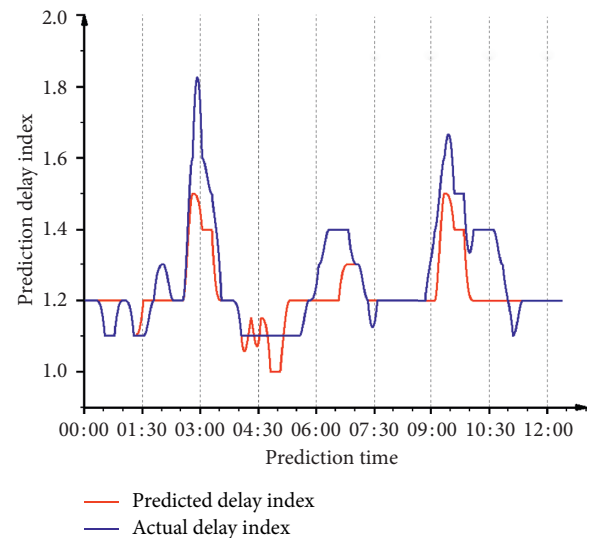
FIGURE 11: Comparison of the predicted and the actual delay index for group f.

accuracy of the proposed method is compared with the LSTM, KNN, SVM, exponential smoothing, and BP neural network. All prediction processes are based on the data used in this paper. Finally, the two representative indicators of prediction accuracy and MSE are used to measure the effectiveness of the forecast.

TABLE 10: The prediction accuracy of groups 1–11.

	1	2	3	4	5	6	7	8	9	10	11
Accuracy (%)	**91.950**	90.525	91.675	90.550	91.030	91.400	91.125	91.075	90.638	90.875	90.200

TABLE 11: The prediction accuracy of groups a–g.

	a	b	c	d	e	f	g
Accuracy (%)	91.850	91.202	89.528	91.600	91.400	**94.178**	91.328
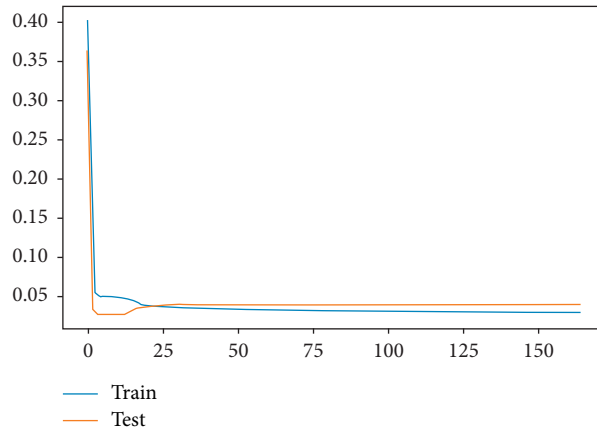
FIGURE 12: LSTM-based traffic forecast.

TABLE 12: SVM and KNN prediction accuracy table.

	SVM			KNN		
Error rate	Accuracy	Recall rate	F1	Accuracy	Recall rate	F1
0.2687	0.7292	0.7332	0.7293	0.8357	0.8409	0.8369

Note. $F1$ represents an indicator of the accuracy of the binary classification model. It also takes into account the accuracy rate and recall rate of the classification model.
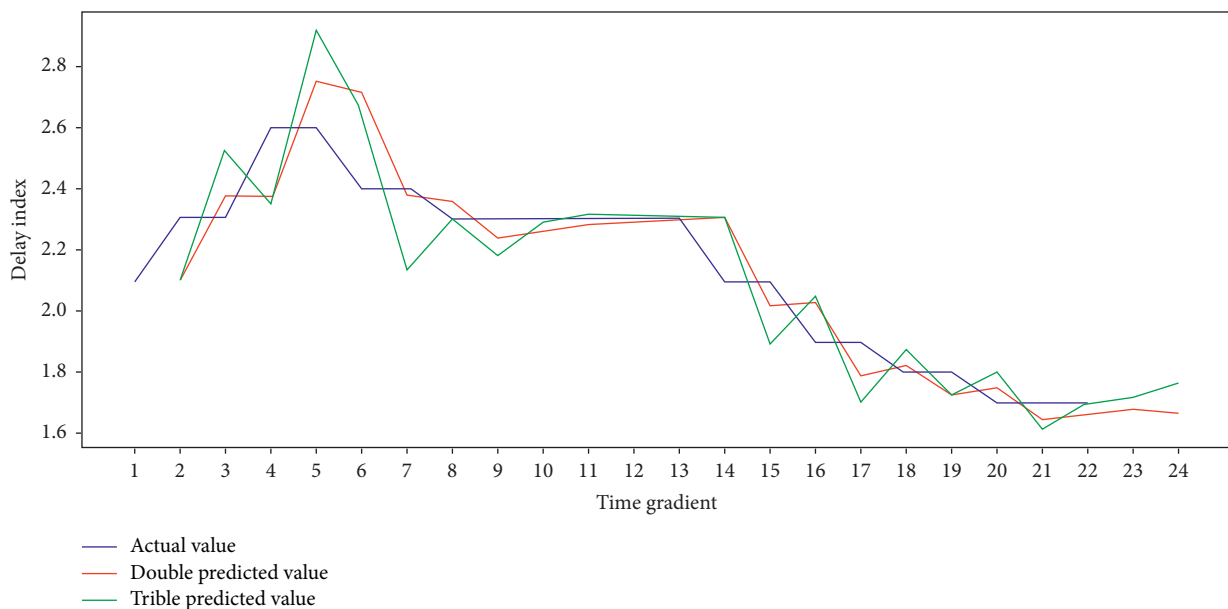
FIGURE 13: Traffic forecast based on the exponential smoothing model.
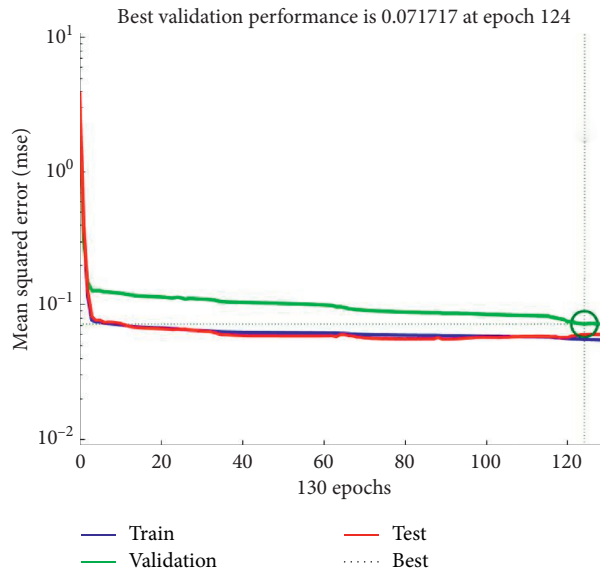
Figure 14: Distribution of mean squared error.
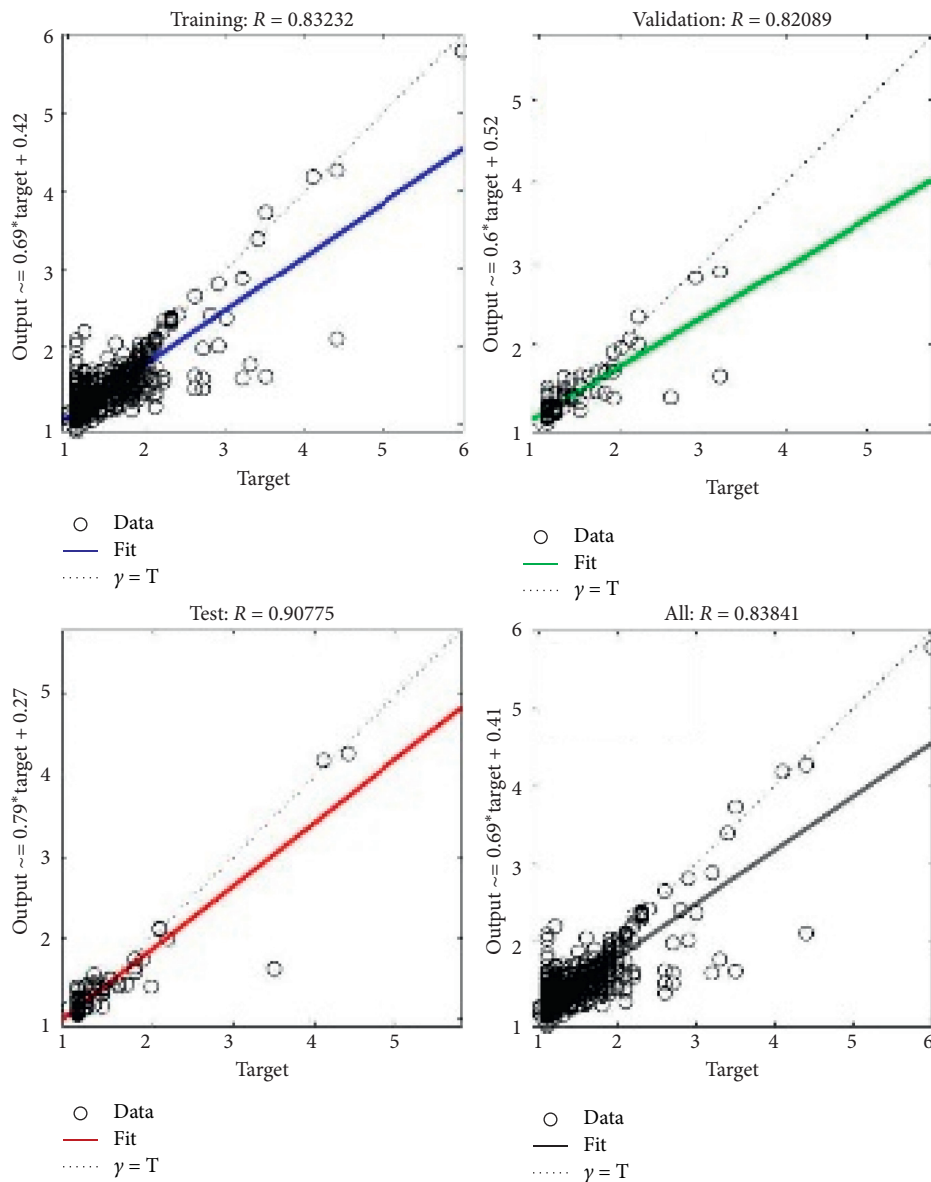
Figure 15: Prediction result graph.

TABLE 13: Comparison table of accuracy of all prediction methods.

	LSTM and Q-Learning	LSTM	KNN	SVM	Exponential smoothing	BP neural network
Prediction accuracy	**94.178%**	88.947%	83.570%	72.920%	82.898%	83.841%
MSE	**0.0226**	0.0595	0.0540	0.0722	0.1710	0.0717

In the LSTM, the data from 0:00 on May 17th to 24:00 on May 24th are used as the training sets, and the data from 0:00 to 12:00 on May 25th are used as the test sets. The prediction results are shown in Figure 12.

In KNN and SVM prediction, the prediction accuracy of the SVM and KNN based on data analysis is shown in Table 12.

In the exponential smoothing forecast, based on the May 20th 0: 00–22 o'clock to predict the traffic state of 23–24 o'clock, the second exponential smoothing and the third exponential smoothing are performed and, finally, compared with the actual delay index. The experimental results show that the fitting curve under quadratic exponential smoothing is better than the prediction result, as shown in Figure 13.

In BP neural network prediction, speed, delay time, travel time, air temperature, and precipitation probability are used as the input matrix, and delay index is used as the output matrix. 90% are used as the training set, 5% are used as the validation set, 5% are used as the prediction set, and 10 hidden layers are used to construct the BP neural network. Training according to the Levenberg–Marquardt algorithm is performed until the best effect is achieved. Finally, the error distribution and prediction results are obtained, as shown in Figure 14 and 15.

It can be seen from Table 13 that the method proposed in this paper is more accurate than LSTM alone. Comparing the proposed method with other representative prediction methods, it is obvious that the prediction effect is better. Therefore, it further confirms the superiority of the method proposed in this paper, which can meet the demand for high efficiency and precision in traffic prediction and has the feasibility of practical application.

7. Conclusions

This paper proposed a short-term traffic flow prediction model for urban roads based on the LSTM and Q-Learning, which are used to solve the problems of low temporal correlation of traffic data, large inventory, poor comprehensive analysis, and slow feedback of prediction results. The analysis results showed that the model has excellent stability and prediction accuracy. Therefore, this model has the feasibility to apply to actual traffic scenarios and to provide accurate information guidance to reduce traffic congestion and accident rates. Moreover, this model could provide substantial method support with the development of active safety.

At the same time, the problem with this model is that the amount of data and data dimensions predicted by this training are not big enough. If there are sufficient data volume and dimensions, it will bring more mature training effects and prediction results. Therefore, the next research goal is to develop more multidimensional research directions based on deep mining of effective traffic data.

In the future, we will focus on exploring more efficient prediction methods based on the research results of this paper. Also, a series of traffic conditions such as future traffic flow, accident trends, and driving behavior trends will be predicted by introducing more relevant data.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] L. Hu, X. T. Hu, J. Wan, M. Lin, and J. Huang, "The injury epidemiology of adult riders in vehicle-two-wheeler crashes in China, Ningbo, 2011–2015," *Journal of Safety Research*, vol. 72, pp. 21–28, 2019.

[2] Y. Y. Guo, P. Liu, Y. Wu, and F. Liu, "Traffic conflict model based on Bayesian multivariate Poisson-lognormal normal distribution," *China Journal of Highway and Transport*, vol. 31, no. 1, pp. 101–109, 2018.

[3] Y. Guo, P. Liu, Y. Wu, and J. Chen, "Evaluating how right-turn treatments affect right-turn-on-red conflicts at signalized intersections," *Journal of Transportation Safety & Security*, vol. 12, no. 3, pp. 419–440, 2020.

[4] Y. Guo, M. Essa, T. Sayed, M. M. Haque, and S. Washington, "A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 96–110, 2019.

[5] L. Hu, J. Ou, J. Huang, Y. Chen, and D. Cao, "A review of research on traffic conflicts based on intelligent vehicles," *IEEE Access*, vol. 8, pp. 24471–24483, 2020.

[6] H. Zhao, Y. Li, W. Hao, S. Peeta, and Y. Wang, "Evaluating the effects of switching period of communication topologies and delays on electric connected vehicles stream with car-following theory," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020.

[7] Y. Li, W. Chen, S. Peeta, and Y. Wang, "Platoon control of connected multi-vehicle systems under V2X communications: design and experiments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 5, pp. 1891–1902, 2020.

[8] J. Tang, Y. Wang, and F. Liu, "Characterizing traffic time series based on complex network theory," *Physica A: Statistical*

Mechanics and Its Applications, vol. 392, no. 18, pp. 4192–4201, 2013.

[9] M. A. Quddus, "Time series count data models: an empirical application to traffic accidents," *Accident Analysis & Prevention*, vol. 40, no. 5, pp. 1732–1741, 2008.

[10] M. Van Der Voort, M. Dougherty, and S. Watson, "Combining Kohonen maps with arima time series models to forecast traffic flow," *Transportation Research Part C: Emerging Technologies*, vol. 4, no. 5, pp. 307–318, 1996.

[11] B. L. Smith, B. M. Williams, and R. Keith Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 4, pp. 303–321, 2002.

[12] G. A. Davis and N. L. Nihan, "Nonparametric regression and short-term freeway traffic forecasting," *Journal of Transportation Engineering*, vol. 117, no. 2, pp. 178–188, 1991.

[13] X. Y. Gong and S. M. Tang, "Integrated traffic flow forecasting and traffic incident detection algorithm based on non-parametric regression," *China Journal of Highway and Transport*, vol. 16, no. 1, pp. 83–87, 2001.

[14] R. E. Turochy and B. D. Pierce, "Relating short-term traffic forecasting to current system state using nonparametric regression," in *Proceedings of the 2004 IEEE Intelligent Transportation Systems Conference*, pp. 239–244, Washington, DC, USA, 2004.

[15] J. J. Tang, Y. W. Wang, W. Hao, F. Liu, H. L. Huang, and Y. H. Wang, "A mixed path size logit-based taxi customer-search model considering spatio-temporal factors in route choice," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 4, pp. 1347–1358, 2019.

[16] C. Qing and W. Hao, "A methodology for measuring and monitoring congested corridors: applications in Manhattan using taxi GPS data," *Journal of Urban Technology*, vol. 24, no. 4, pp. 1–17, 2018.

[17] Z. W. Wang, J. Yu, W. Hao, and J. J. Tang, "Two-step co-ordinated optimization model of mixed demand responsive feeder transit," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 3, pp. 1–12, 2020.

[18] W. Hao, Z. Zhang, Z. Gao, K. Yi, L. Liu, and J. Wang, "Research on mandatory lane-changing behavior in highway weaving sections," *Journal of Advanced Transportation*, vol. 2020, Article ID 3754062, 9 pages, 2020.

[19] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 3, pp. 211–234, 2005.

[20] X. Y. Gong and S. M. Tang, "Integrated traffic flow forecasting and traffic incident detection algorithm based on non-parametric regression," *China Journal of Highway and Transport*, vol. 16, no. 1, pp. 82–86, 2003.

[21] Z. S. Yang, Y. Wang, and Q. Guan, "Short-term traffic flow prediction method based on SVM," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 36, no. 6, pp. 881–884, 2006.

[22] Z. S. Yao, C. F. Shao, and Z. H. Xiong, "Short-term traffic volumes forecasting of road network based on principal component analysis and support vector machine," *Journal of Jilin University*, vol. 38, no. 1, pp. 48–52, 2008.

[23] Y.-S. Jeong, Y.-J. Byon, M. M. Castro-Neto, and S. M. Easa, "Supervised weighting-online learning algorithm for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1700–1707, 2013.

[24] B. Yu, S. H. Wu, and M. H. Wang, "K-nearest neighbor model of short-term traffic flow forecast," *Journal of Traffic and Transportation Engineering*, vol. 12, no. 2, pp. 105–111, 2012.

[25] H. Chang, Y. Lee, B. Yoon, and S. Baek, "Dynamic near-term traffic flow prediction: system-oriented approach based on past experiences," *IET Intelligent Transport Systems*, vol. 6, no. 3, pp. 292–305, 2012.

[26] H. H. Xie, X. H. Dai, and Y. Qi, "Improved K-nearest neighbor algorithm for short-term traffic flow forecasting," *Journal of Traffic and Transportation*, vol. 14, no. 3, pp. 87–94, 2014.

[27] Q. Bing, B. Gong, Z. Yang, Q. Shang, and X. Zhou, "Short-term traffic flow local prediction based on combined Kernel function relevance vector machine model," *Mathematical Problems in Engineering*, vol. 2015, Article ID 154703, 9 pages, 2015.

[28] Q. Shang, Z. S. Yang, and Z. L. Li, "Short-term traffic flow prediction based on phase space reconstruction and RELM," *Journal of South China University of Technology (Natural Science Edition)*, vol. 44, no. 4, pp. 109–114, 2016.

[29] B. Moghimi, A. Safikhani, C. Kamga, W. Hao, and J. Q. Ma, "Short-term prediction of signal cycle on an arterial with actuated-uncoordinated control using sparse time series models," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 8, pp. 2976–2985, 2018.

[30] B. Moghimi, A. Safikhani, C. Kamga, and W. Hao, "Cycle-length prediction in actuated traffic-signal control using ARIMA model," *Journal of Computing in Civil Engineering*, vol. 32, no. 2, pp. 1–10, 2018.

[31] J. Ma, X. Li, F. Zhou, and W. Hao, "Designing optimal autonomous vehicle sharing and reservation systems: a linear programming approach," *Transportation Research Part C: Emerging Technologies*, vol. 84, pp. 124–141, 2017.

[32] K. Saito, M. Kimura, K. Ohara, and H. Motoda, "Efficient discovery of influential nodes for SIS models in social networks," *Knowledge and Information Systems*, vol. 30, no. 3, pp. 613–635, 2012.

[33] F. Wang, H. Y. Wang, K. Xu, J. H. Wu, and X. H. Jia, "Characterizing information diffusion in online social networks with linear diffusive model," in *Proceedings of the IEEE 33rd International Conference on Distributed Computing Systems*, pp. 307–316, Philadelphia, PA, USA, 2013.

[34] T. Kuremoto, S. Kimura, K. Kobayashi, and M. Obayashi, "Time series forecasting using a deep belief network with restricted Boltzmann machines," *Neurocomputing*, vol. 137, pp. 47–56, 2014.

[35] Y. Lv, Y. Duan, W. Kang, Z. X. Li, and F. Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.

[36] H. Z. Zhao, D. H. Sun, M. Zhao, and S. L. Cheng, "A multi-classification method of improved SVM-based information fusion for traffic parameters forecasting," *Promet—Traffic & Transportation*, vol. 28, no. 2, pp. 17–24, 2016.

[37] T. Djukic, J. W. C. Van Lint, and S. P. Hoogendoorn, "Application of principal component analysis to predict dynamic origin-destination matrices," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2283, no. 1, pp. 81–89, 2012.

[38] N. Mitrovic, M. T. Asif, U. Rasheed, J. Dauwels, and P. Jaillet, "CUR decomposition for compression and compressed sensing of large-scale traffic data," in *Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems*, pp. 1475–1480, The Hague, Netherlands, 2013.

[39] M. T. Asif, S. Kannan, J. Dauwels, and P. Jaillet, "Data compression techniques for urban traffic data," in *Proceedings of the IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS)*, pp. 44–49, Singapore, 2013.

[40] Y. Y. Chen, Y. S. Lv, Z. J. Li, and F. Y. Wang, "Long short-term memory model for traffic congestion prediction with online open data," in *Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 132–137, Rio de Janeiro, Brazil, 2016.

[41] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68–75, 2017.

[42] X. Luo, D. Li, Y. Yang, and S. Zhang, "Spatiotemporal traffic flow prediction with KNN and LSTM," *Journal of Advanced Transportation*, vol. 2019, Article ID 4145353, 10 pages, 2019.

[43] F. Chen, H. R. Peng, X. X. Ma, J. Y. Liang, W. Hao, and X. D. Pan, "Examining the safety of trucks under crosswind at bridge-tunnel section: a driving simulator study," *Tunnelling and Underground Space Technology*, vol. 92, 2019.

[44] A. Miglani and N. Kumar, "Deep learning models for traffic flow prediction in autonomous vehicles: a review, solutions and challenges," *Vehicular Communications*, vol. 20, pp. 10–20, 2019.

[45] B. Ji and E. J. Hong, "Deep-learning-based real-time road traffic prediction using long-term evolution access data," *Sensors*, vol. 19, no. 23, pp. 1–10, 2019.

[46] Y. Zhang, T. Cheng, Y. B. Ren, and K. Xie, "A novel residual graph convolution deep learning model for short-term network-based traffic forecasting," *International Journal of Geographical Information Science*, vol. 34, no. 5, pp. 969–995, 2020.

[47] M. Abdollahi, T. Khaleghi, and K. Yang, "An integrated feature learning approach using deep learning for travel time prediction," *Expert Systems with Applications*, vol. 139, pp. 1–10, 2019.

WILEY | Hindawi

Research Article
Customized Bus Route Optimization with the Real-Time Data

Kai Huang,[1,2,3] Lin Xu,[4] Yao Chen [ID],[5] Qixiu Cheng [ID],[1] and Kun An[2]

[1]*Jiangsu Key Laboratory of Urban ITS Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Suzhou, China*
[2]*College of Transportation Engineering, Tongji University, Shanghai, China*
[3]*Department of Civil Engineering, Monash University, Melbourne, Australia*
[4]*Jiangsu Vocational Institute of Architectural Technology, Xuzhou, China*
[5]*MOE Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing, China*

Correspondence should be addressed to Qixiu Cheng; qixiu.cheng@seu.edu.cn

This paper investigates the real-time customized bus (CB) route optimization problem, which aims to maximize the service rate for clients and profits for operators. The on-road bus has a flexible route, which can be updated based on the real-time data and route optimization solutions. A two-phase framework is established. In phase 1, the vehicle-related data including existing route and schedule, client-related data involving pick-up/drop-off location, and time windows are collected once receiving a new CB request. The second phase optimizes the bus route by establishing three nonlinear programming models under the given data from phase 1. A concept of profit difference is introduced to decide the served demand. To improve computation efficiency, a real-time search algorithm is proposed that the neighboring buses are tested one by one. Finally, a numerical study based on Sioux Falls network reveals the effectiveness of the proposed methodology. The results indicate that the real-time route optimization can be achieved within the computation time of 0.17–0.38 seconds.

1. Introduction

In recent years, a series of traffic problems happen due to traffic congestion in an urban area, such as increasing on-road travel time of citizens during daily commuting. Encouraging public transport, like buses, is one of the main methods to address such a problem [1, 2]. Multiple clients can share a vehicle, which only occupy very limited traffic resources. The travel fee is obviously lower than the taxi and private car. However, the disadvantage of taking the bus is the low level of service (LOS) caused by the long travel time. Bus clients have to walk for a long distance to the fixed station from their desired origins, home, or workplace, which would largely harm their enthusiasm. Hence, the customized bus (CB) is proposed and becomes popular in many cities around the world. It has a flexible route to pick-up waiting clients in their origins. CB plays the role of a bridge between the traditional buses and taxi/private car. CB

provides the ride-hailing service that clients can make travel requests by using mobility apps before departure. Clients can publish their desired pick-up/drop-off location and time windows. For the CB companies, named CB operators, they can dynamically change the route and schedule based on the real-time travel request.

CB system can be divided into two main categories based on clients' OD distribution: many-to-one mode, that is, clients have different origins but the same destination; many-to-many mode services, that is, clients have different origins and destinations. The former can meet the demand of clients who have the same trip purpose. Taking the passengers whose destination is the airport as an example, they can share a CB by publishing request online days or hours earlier [3]. Many existing studies have explored the many-to-one CB service [4]. When considering the return trip, a concept named as many-to-one-to-many is proposed by Tarantilis et al. [5]. It can be seen as two independent many-

to-one trips when optimizing the bus route. Moreover, to provide more flexible service, many-to-many service is raised that clients are allowed to share a bus if they have an overlapping route [6]. Such a mode has been widely conducted in the shared mobilities, like the taxi sharing and carpooling [7]. In terms of the multiple pick-up and drop-off stations, it makes the system optimization more complex [8].

Another difficulty in the CB system service, especially the many-to-many service, is the vehicle routing problem (VRP) [9]. The order of station visiting can largely affect the LOS of in-vehicle and waiting clients. A reasonable route can ensure the service reliability. The existing studies normally establish optimization models to address the CB VRP, like the schedule coordination submodel for the feeder bus [10] and integer programming model for the school bus [11]. In the objective function, the existing studies normally either maximize the service rate or minimize the operation costs [12–14]. However, it is challenging to find a balance between the two aspects. The low service rate results in the potential demand loss, while the high operation cost makes the CB system unsustainable [15–18].

The last challenge comes from the real-time demand response. Studies show that only limited papers explore the real-time CB system optimization [19]. With the use of mobility apps, clients are allowed to make a request once he or she decides to take buses. It brings a huge challenge for the CB operator to collect the data, optimize bus routes, and then update the schedule in a short time. Hence, reducing the computation time is essential when providing the real-time decisions.

This paper aims to maximize the service rate of CB clients and profits of CB operators. A real-time two-phase research methodology is built. In phase 1, real-time data are collected once a new CB request is proposed. It mainly includes the vehicle-related data with existing route and schedule and client-related data with pick-up/drop-off location and time windows. And then, the collected data are set as the inputs of phase 2, where three CB route optimization models are established. Considering the location of the new request, the many-to-many CB mode can be divided into three categories: only new pick-up service but drop-off location in the existing route, only new drop-off service but pick-up location in the existing route, and both new pick-up and drop-off service. The profit difference is introduced to decide the served demand. To handle the optimization problem, a real-time search algorithm is proposed that the neighboring buses are tested one by one.

In the next section, the research methodology is given that including the real-time data collection in phase 1, the optimization model, and solution approach in phase 2. Section 3 presents the application to the Sioux Falls network. Conclusions and future work directions are presented in the last section.

2. Research Methodology

2.1. Problem Setting.
We consider a real-time CB route optimization problem that the operators update the bus route once a new request, namely, the latest request, is claimed. The objective of this paper is to find a vehicle to meet the latest request with the objective of maximizing service rate for clients and profit for CB operators. The study region consists of $|I|$ bus stations with a parking station denoted as $i \in I$. The current location is taken as the first station. There are $|V|$ vehicles in the CB system with a bus denoted as $v \in V$. The bus v has a fixed route to drop off the in-vehicle clients and pick up waiting clients. The bus will be used to serve the latest client from station $i \in I$ to station $j \in I$ where $i \neq j$. It is defined by decision variables Y_i and X_{ji}. Y_i equals 1 if vehicle $v \in V$ arrives at station $i \in I$ and 0 otherwise. X_{ji} equals 1 if vehicle departs from station $j \in I$ to station $i \in I$, where $i \neq j$ and 0 otherwise. Similarly, X_{ij} represents the binary decision of vehicle travel from station $i \in I$ to station $j \in I$. The fixed time window between bus arriving at and departing from each station is given. Let tep_i be the desired earliest pick-up time of clients at bus station $i \in I$, tlp_i be the latest pick-up time of clients at bus station $i \in I$, and tld_i be the latest drop-off time of clients at bus station $i \in I$. Clients who have the similar time windows and the same stations are grouped into one trip.

In practice, the pick-up and drop-off location of the latest request might not always exist in the given route. Based on the station location, we divide the CB route optimization problem into three types, as shown in Figure 1. Figure 1(a) shows the CB system with a single external pick-up station. The client's origin location is not in the existing route, while her/his destination location is the bus route. In contrast, Figure 1(b) represents the CB system with a single external drop-off station. Figure 1(c) illustrates the CB system with external pick-up and drop-off stations. Both the client's pick-up and drop-off stations are not in the existing bus route. The route optimization is to decide serving the latest request or not by updating the bus route. For an empty bus in the parking station, the trip of latest request can be seen as the optimized route.

2.2. Notation.
The notation used throughout the paper is listed below:

Sets

I: $\{i\}$, $\{j\}$: set of bus stations
V: $\{v\}$: set of buses

Parameters

c: travel cost of bus per time step
p: payment of the latest CB request
tt_{ij}: travel time from bus station $i \in I$ to bus station $j \in I$, where $i \neq j$
tep_i: desired earliest pick-up time of client at bus station $i \in I$
tlp_i: latest pick-up time of client at bus station $i \in I$
tld_i: latest drop-off time of client at bus station $i \in I$

Decision variables

X_{ij}: binary variable; it equals 1, and if vehicle travels from station $i \in I$ to station $j \in I$, where $i \neq j$, it equals 1 and 0 otherwise
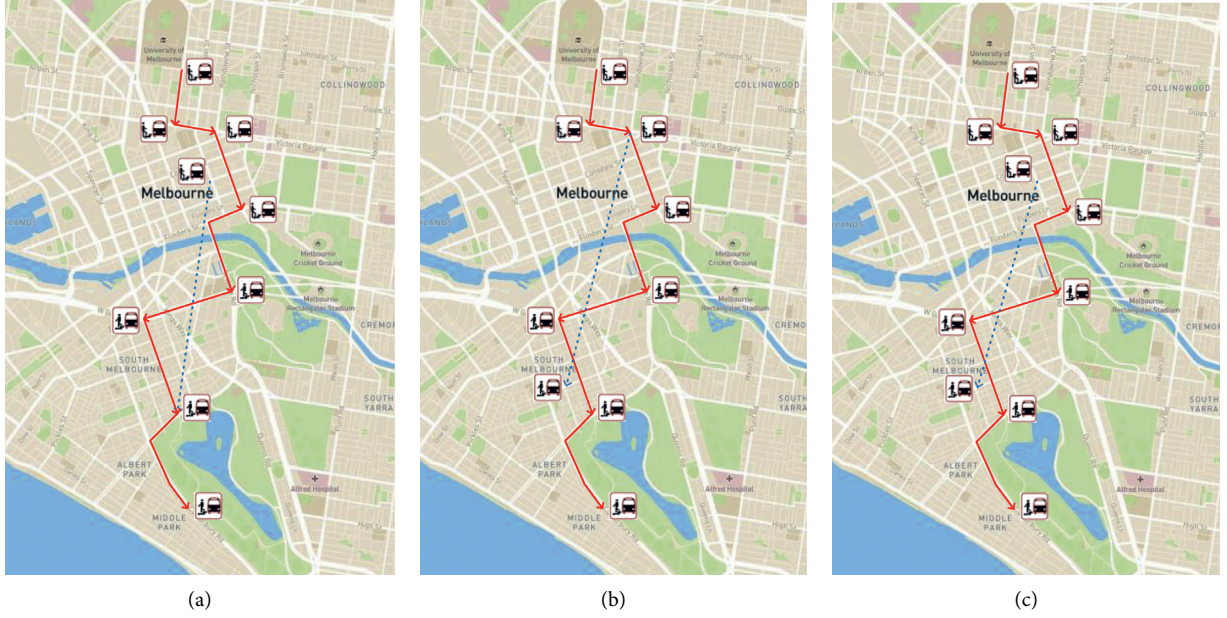
FIGURE 1: The customized bus system.

Y_i: binary variable; it equals 1, and if vehicle arrives at station $i \in I$, it equals 1 and 0 otherwise

Auxiliary variables

ta_i: arriving time of vehicle at bus station $i \in I$
td_i: departure time of vehicle at bus station $i \in I$
μ_i, ν_i: binary variables involved in the linearization

2.3. Phase 1: Real-Time Data Collection.

In phase 1, the CB client claims the real-time CB request that includes pick-up location, drop-off location, time window of the desired earliest pick-up time, latest pick-up time, and latest drop-off time. Once such a request is published, the CB operator will search neighboring buses and identify vehicle location, existing route and schedule, in-vehicle and waiting clients' location, and time window constraints. The neighboring search radius is the distance between the bus current location and latest clients' pick-up location. The vehicle- and client-related data can be set as inputs of the optimization model in phase 2. The real-time CB route optimization can be fulfilled by rolling the bus search in phase 1 and optimization in phase 2. The profit difference among all tested buses decides the selected bus.

2.4. Phase 2: Route Optimization Model

2.4.1. M-1: CB System with Single External Pick-Up Station.
The model with a single external pick-up station is formulated as follows. We set the picking up location of the latest CB request as station $|I|$:

$$\textbf{P1-1} \max_{X,Y} \phi = Y_{|I|} \tag{1}$$

subject to

$$\sum_{i \in I} X_{i1} = 0, \tag{2}$$

$$\sum_{j \in I} X_{1j} = 1, \tag{3}$$

$$X_{ii} = 0, \quad \forall i \in I, \tag{4}$$

$$\sum_{j \in I} X_{ij} \leq 1, \quad \forall i \in I, \tag{5}$$

$$\sum_{i \in I} X_{ij} = 1, \quad \forall j = 2, \ldots, |I| - 1, \tag{6}$$

$$\sum_{j \in I} X_{j|I|} = \sum_{j \in I} X_{|I|j}, \tag{7}$$

$$Y_i = \max \left(\sum_{j \in I} X_{ij}, \sum_{j \in I} X_{ji} \right), \quad \forall i = I, \tag{8}$$

$$ta_1 = 0, \tag{9}$$

$$td_1 = 0, \tag{10}$$

$$ta_i \leq \min \left(Y_i tlp_i, Y_i tld_i \right), \quad \forall i = 2, \ldots, |I|, \tag{11}$$

$$td_i \geq \max \left(Y_i ta_i, Y_i tep_i \right), \quad \forall i = 2, \ldots, |I|, \tag{12}$$

$$ta_j = \sum_{i \in I} X_{ij} td_i + \sum_{i \in I} X_{ij} tt_{ij}, \quad \forall j = 2, \ldots, |I|, \tag{13}$$

$$X_{ij}, Y_i \in \{0,1\}, ta_i, td_i \ge 0, \quad \forall i \in I, j \in I. \tag{14}$$

The objective function (1) aims to maximize the service rate of the latest CB request. Constraint (2) defines the starting station of a bus, where the vehicle is found when the latest CB request is proposed. Constraint (3) asks the bus must first leave the starting station and then serve other demand in the next trips. Only one trip between the starting station and other stations is approved. Constraint (4) represents that a bus cannot travel within its own station. Constraint (5) ensures the no more than one trip leaving one station is allowed. If one station is the ending station of this CB service, the trip departing from it would not exist. The term $\sum_{j\in I} X_{ij}$ represents all the possible trips leaving station i. Constraint (6) asks the in-vehicle and waiting clients, whose request with the pick-up or drop-off station $j = 2, \ldots, |I| - 1$ has to be satisfied. The term $\sum_{j\in I} X_{ji}$ represents all the probable trips arriving at station i. Constraint (7) indicates that the bus must first arrive at and then leave the station, where the latest request happens if the client is served. Otherwise, no bus is arranged to visit this station. Constraint (8) identifies whether the request in one station is served or not based on the bus trip. Constraints (9) and (10) state that the arriving time and departing time are set as 0 in the starting station. Constraint (11) ensures that the bus arrives at one station no later than the waiting clients' latest pick-up time and the in-vehicle clients' latest drop-off time. Constraint (12) ensures that the bus leaves one bus station no earlier than the waiting clients' earliest pick-up time and bus arriving time. Constraint (13) calculates the vehicle arriving time. The term $\sum_{i\in I} X_{ij} td_i$ and $\sum_{i\in I} X_{ij} tt_{ij}$ represent the departing time and travel time of the bus at the station i. Constraint (14) specifies the domain of the decision variables.

Based on the optimized route in **P1-1**, we calculate the increased profits of the latest CB request φ and total travel time cost γ_1:

$$\varphi = Y_{|I|} p, \tag{15}$$

$$\gamma_1 = \sum_{i\in I} \sum_{j\in J} X_{ij} tt_{ij} c, \tag{16}$$

$$\textbf{P1-2} \min_{X,Y} \phi = Y_{|I|}. \tag{17}$$

Subject to Constraints (2)–(14).

The objective function (17) minimizes the service rate of the latest CB request. It shares the same constraints with **P1-1**. Based on the optimized route in **P1-2**, we calculate the total travel time cost γ_2 with equation (16).

The profit difference is obtained as follows:

$$\textbf{P1-3} \phi = \varphi - \gamma_1 + \gamma_2. \tag{18}$$

Equation (18) calculates the profit difference between **P1-1** and **P1-2**. If the latest CB request in **P1-1** is not served, ϕ would be equal to 0. Otherwise, it will be a positive or

negative value that depends on the increased income and operation cost caused by servicing the latest CB request.

2.4.2. M-2: CB System with Single External Drop-Off Station. The model with a single external drop-off station is formulated as follows. We set the dropping off location of the latest CB request as station $|I|$:

$$\textbf{P2-1} \max_{X,Y} \phi = Y_{|I|}. \tag{19}$$

Constraints (2)–(6) and (9)–(14), plus:

$$Y_i = \max\left(\sum_{j\in I} X_{ij}, \sum_{j\in I} X_{ji}\right), \quad \forall i = |I| - 1, \tag{20}$$

$$Y_{|I|} = \sum_{j\in I} X_{j|I|}. \tag{21}$$

Constraint (20) identifies whether the request in one station of the existing route is served or not based on the bus trip. Constraint (21) identifies whether the bus arrives at the drop-off station of the latest request or not based on the arriving trip. As this station can be set as the ending station, the departing trips leaving this station might not exist.

Based on equations (15) and (16), we calculate the increased profits of the latest CB request φ and the total travel time cost γ_1 in **P2-1**:

$$\textbf{P2-2} \min_{X,Y} \phi = Y_{|I|}. \tag{22}$$

Subject to Constraints (2)–(6), (9)–(14), (20), and (21). Based on equation (16), we calculate the total travel time cost γ_2 in **P2-2**:

$$\textbf{P2-3} \phi = \varphi - \gamma_1 + \gamma_2. \tag{23}$$

The profit difference between **P1-1** and **P1-2** is obtained based on equation (23).

2.4.3. M-3: CB System with Both External Pick-Up and Drop-Off Stations. The model with both external pick-up and drop-off stations is formulated as follows. We set the picking up location of the latest CB request as station $|I| - 1$ and dropping off location as station $|I|$:

$$\textbf{P3-1} \max_{X,Y} \phi = Y_{|I|-1}. \tag{24}$$

Subject to Constraints (2)–(5), (9)–(14), (20), and (21), plus

$$\sum_{i\in I} X_{ij} = 1, \quad \forall j = 2, \ldots, |I| - 2, \tag{25}$$

$$\sum_{j\in I} X_{j|I|-1} = \sum_{j\in I} X_{|I|-1j}, \tag{26}$$

$$Y_{|I|} = Y_{|I|-1}. \tag{27}$$

Constraint (25) asks the existing requests, picking up or dropping off at station $j = 2, \ldots, |I| - 2$, in the existing route must be served. Constraint (26) indicates that the bus must

first arrive at and then leave from the station where the latest request happens if this request is met. Constraint (27) represents that the bus has to visit pick-up and drop-off stations if the latest request is served. Otherwise, the bus will not arrive at both stations.

Based on equations (15) and (16), we calculate the increased profits of the latest CB request φ and total travel time cost γ_1 in **P3-1**:

$$\textbf{P3-2} \min_{X,Y} \phi = Y_{|I|-1}. \tag{28}$$

Subject to Constraints (2)–(5), (9)–(14), (20), (21), and (25)–(27).

Based on equation (16), we calculate the total travel time cost γ_2 in **P3-2**:

$$\textbf{P3-3} \ \phi = \varphi - \gamma_1 + \gamma_2. \tag{29}$$

The profit difference between **P3-1** and **P3-2** is obtained based on equation (29).

2.4.4. Solution Approach

(1). Linearization. The nonlinear constraint (8) incurs a great computation challenge. A series of linearization methods are proposed as follows: constraint (8) is replaced by constraints (31)–(36):

$$Y_i \geq \sum_{j \in I} X_{ij}, \quad \forall i = I, \tag{30}$$

$$Y_i \geq \sum_{j \in I} X_{ji}, \quad \forall i = I, \tag{31}$$

$$Y_i - M(1 - \mu_i) \leq \sum_{j \in I} X_{ij}, \quad \forall i = I, \tag{32}$$

$$Y_i - M(1 - \nu_i) \leq \sum_{j \in I} X_{ji}, \quad \forall i = I, \tag{33}$$

$$\mu_i + \nu_i \geq 1, \quad \forall i = I, \tag{34}$$

$$\mu_i, \nu_i \in \{0, 1\}, \quad \forall i = I, \tag{35}$$

when $\mu_{it} = 1$ and $\nu_{it} = 0$. Constraint (32) reduces to $Y_i \leq \sum_{j \in I} X_{ij}$, and constraint (33) is redundant. Together with constraint (30), we have $Y_i = \sum_{j \in I} X_{ij}$, and constraint (31) reduces to $\sum_{j \in I} X_{ij} \geq \sum_{j \in I} X_{ji}$. When $\mu_{it} = 0$ and $\nu_{it} = 1$, constraint (33) reduces to $Y_i \leq \sum_{j \in I} X_{ji}$ and constraint (32) is redundant. Together with constraint (31), we have $Y_i = \sum_{j \in I} X_{ji}$. When $\mu_{it} = 1$ and $\nu_{it} = 1$, we have $Y_i = \sum_{j \in I} X_{ij}$ and $Y_i = \sum_{j \in I} X_{ji}$.

(2). Search Algorithm. We develop a real-time search algorithm to optimize the CB system. It can reduce the computation burden by searching the bus one by one. In this way, the maximum profits for CB operators can be obtained via rolling all the found buses. The following pseudocode shows such an algorithm:

Step 1: choose the real-time traffic network at time instant of the latest request happening, and set the initial profit difference ϕ as 0.

Step 2: search the neighboring buses V of the latest CB request with a given search radius and number of required seats.

Step 3: choose the vehicle v among all the found buses V, and check the existing route of this vehicle.

Step 3.1: if both the pick-up location and drop-off location of the latest request are in the existing route, go to step 4; otherwise, proceed to step 3.2.
Step 3.2: if only the pick-up location is in the existing route, go to step 5; otherwise, proceed to step 3.3.
Step 3.3: if only the drop-off location is in the existing route, go to step 6; otherwise, proceed to step 3.4.
Step 3.4: if both the pick-up location and drop-off location of the latest request are not in the existing route, go to step 7.

Step 4: check the time window constraints. If the selected bus can serve the latest request, arrange the client to it, stop. Otherwise, proceed to step 3.

Step 5: optimize models **P1-1** and **P1-2** and then calculate the profit difference by equation (18). Save the result as ϕ_v.

Step 5.1: compare the ϕ_v and ϕ. If $\phi_v > \phi$, update the optional bus as v.
Step 5.2: if $v = |V|$, stop. Otherwise, set $v = v + 1$ and go to step 3.

Step 6: optimize models **P2-1** and **P2-2** and then calculate the profit difference by equation (23). Save the result as ϕ_v.

Step 6.1: compare the ϕ_v and ϕ. If $\phi_v > \phi$, update the optional bus as v.
Step 6.2: if $v = |V|$, stop. Otherwise, set $v = v + 1$ and go to step 3.

Step 7: optimize models **P3-1** and **P3-2** and then calculate the profit difference by equation (29). Save the result as ϕ_v.

Step 7.1: compare the ϕ_v and ϕ. If $\phi_v > \phi$, update the optional bus as v.
Step 7.2: if $v = |V|$, stop. Otherwise, set $v = v + 1$ and go to step 3.

Moreover, the parallel computing technique can be used to save computation time. The route optimization of each bus in phase 2 is independent that can be carried out at the same time with multiprocessing modules. In this way, the total computation time can be reduced to $1/|V|$.

3. Case Study

3.1. Setting up the Case Study. The classical Sioux Falls network is used to demonstrate the applicability of the proposed models and algorithms. As shown in Figure 2, 24 notes and 38 bidirectional links exist in the network. We
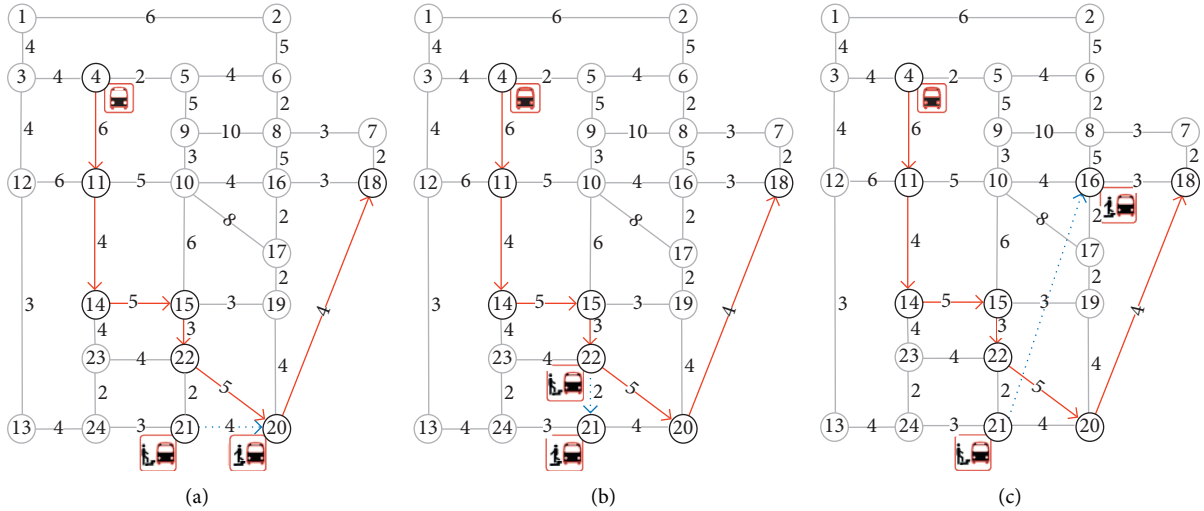
Figure 2: The customized bus system scenarios.

carry out the real-time optimization based on the given data including the client's demand, bus's current location, route, and schedule. We assume that one bus with known routes is already found by the neighboring search, which is providing CB service at the time instant of one new request claiming. Even M-3 can be seen as a general model to optimization all three scenarios, we still use three independent models in this paper. It is to reduce the computation time as required by the real-time requests. An initial test reveals that the total computation time can be reduced to about 50% by using M-1 and M-2 to replace M-3 in special scenarios. Three scenarios are given in this case study. Figure 2(a) represents the M-1 CB system with a single external pick-up station. One client makes a request by starting from station #21 to station #20. Figure 2(b) shows the M-2 CB system with a single external drop-off station. The client departs from station #22 to station #21. Figure 2(c) illustrates the M-3 CB system with both external pick-up and drop-off stations. Station #21 and station #16 are the origin and destination of the latest request.

The client's CB trip is defined by the OD location and pick-up/drop-off time windows at each station in Table 1. The CB trips are not necessary to be tracked in the optimization model, which can save the computation burden by reducing the number of decision variables.

3.2. Optimization Results.
The proposed CB route optimization model is solved by Python calling GUROBI 7.0.2 solver on an i7 processor @2.40 GHz, 8 GB RAM computer with a Windows 7 64 bit operating system. The total computation time is 0.17–0.38 seconds. Table 2 shows the optimized route.

As shown in M-1 of Table 2, we can see that the CB request happening at location #21 can be served that does not break the existing service planning. The in-vehicle and waiting clients can be served within their desired pick-up/drop-off time windows. A new picking up location #21 is put

into the CB route, and the optimized route has been changed to #4-#11-#15-#22-#21-#20-#18. We then made a sensitivity analysis on the time window constraints. We assume that both the earliest and latest drop-off time in the latest request are later than 20 minutes. The results show that there is a time window conflict for the bus to service clients in stations #22 and #21 simultaneously. The bus will follow the old route to serve the existing in-vehicle and waiting clients. In M-2 of Table 2, the optimized route is #4-#11-#15-#22-#21-#20-#18 by adding a new drop-off station #21. The CB system can serve the latest client by taking a detour and does not have a large travel delay. All the clients' time window constraints can be ensured. In M-3 of Table 2, the drop-off location of the latest client is set as the ending station of this CB system. The total operation time rises to 30 minutes. The optimized route is #4-#11-#15-#22-#21-#20-#18-#16. Moreover, the sensitivity analysis in M-2 and M-3 is carried out by using a similar method in M-1. The results reveal that the latest CB request satisfying is largely dependent on the time window constraints.

3.3. Remarks.
In this study, a two-phase research methodology is proposed. Considering the limited techniques to data collection, we assume that the vehicle- and client-related data are known. The CB service route is optimized based on the given data. Although we simplify this phase 1 of data collection, the complete CB route dynamic optimization has been conducted. In addition, the proposed neighboring buses search algorithm can fulfill the real-time CB route optimization. We have recognized the system optimal results cannot be obtained in this study [20], but the computation time and burden can be reduced by testing the bus one by one. In practice, it brings a huge challenge for the existing studies to update the bus route in a short time when optimizing the CB system with multiple vehicles and routes. This paper provides a feasible method to address large-scale many-to-many CB route optimization in an urban area.

TABLE 1: CB request with given OD location and time windows.

Model	Client ID	Origin			Destination	
		Location	tep_i (minutes)	tlp_i (minutes)	Location	tld_i
M-1	1	#4	—	—	#14	11
	2	#4	—	—	#22	20
	3	#11	5	7	#20	26
	4	#14	11	12	#18	30
	5	#15	14	17	#18	30
	6	#21	22	24	#20	26
M-2	1	#4	—	—	#14	11
	2	#4	—	—	#22	20
	3	#11	5	7	#20	26
	4	#14	11	12	#18	30
	5	#15	14	17	#18	30
	6	#22	17	19	#21	22
M-3	1	#4	—	—	#14	11
	2	#4	—	—	#22	20
	3	#11	5	7	#20	26
	4	#14	11	12	#18	30
	5	#15	14	17	#18	30
	6	#21	22	24	#16	33

TABLE 2: Results of CB optimization models.

Mode	Origin			Destination		
	Location	Departing time (minutes)	Pick-up client ID	Location	Arriving time (minutes)	Drop-off client ID
M-1	#4	—	1, 2	#11	—	—
	#11	6	3	#14	6	1
	#14	10	4	#15	11	—
	#15	16	5	#22	16	2
	#22	19	—	#21	19	—
	#21	21	6	#20	22	3, 6
	#20	26	—	#18	26	4, 5
	#18	30	—	—	—	—
M-2	#4	0	1, 2	#11	0	—
	#11	6	3	#14	6	1
	#14	10	4	#15	11	—
	#15	16	5	#22	16	2
	#22	19	6	#21	19	6
	#21	21	—	#20	21	3
	#20	25	—	#18	25	4, 5
	#18	29	—	—	—	—
M-3	#4	0	1, 2	#11	0	—
	#11	6	3	#14	6	1
	#14	10	4	#15	11	—
	#15	16	5	#22	16	2
	#22	19	—	#21	19	
	#21	21	6	#20	22	3
	#20	26	—	#18	26	4, 5
	#18	30	—	#16	30	6
	#16	33	—	—	—	—

4. Conclusion

This paper studies the CB route optimization with real-time data. A two-phase research methodology is established including the data collection and route optimization. In phase 1, the vehicle- and client-related data are real-time collected and then is set as the input to phase 2. Three nonlinear programming models are established to address the many-to-many CB optimization problem. A concept of profit difference is proposed that decides the served requests. A customized solution approach including a linearization method and search algorithm is proposed. The case study based on the Sioux Falls network is carried out to verify the proposed methodology [21–28].

The results show that the proposed optimization model can address the CB request in 0.17–0.38 seconds. It indicates that the real-time CB route optimization is fulfilled by searching the neighboring buses one by one, conducting route optimization, and then comparing the profit difference. The optimization solutions represent that CB request can be satisfied if the bus takes a detour and ensures the existing clients' pick-up/drop-off location and time window constraints simultaneously. Moreover, the sensitivity analysis also verifies the reliability of the proposed models.

Although this paper conducts the CB route optimization, the vehicle- and client-related data are given in the case study. In the future study, the rolling horizon method based on the real-time data can be explored. Furthermore, to save the computation time and burden, we introduce the profit difference. However, the global optimal solutions might not be obtained. Another possible extension is on global optimization of CB route optimization problem.

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 regarding the publication of this article.

Acknowledgments

References

[1] J. Chen, Z. Liu, S. Zhu, and W. Wang, "Design of limited-stop bus service with capacity constraint and stochastic travel time," *Transportation Research Part E: Logistics and Transportation Review*, vol. 83, pp. 1–15, 2015.

[2] Z. Liu, Y. Yan, X. Qu, and Y. Zhang, "Bus stop-skipping scheme with random travel time," *Transportation Research Part C: Emerging Technologies*, vol. 35, no. 9, pp. 46–56, 2013.

[3] J. Park and B.-I. Kim, "The school bus routing problem: a review," *European Journal of Operational Research*, vol. 202, no. 2, pp. 311–319, 2010.

[4] S. N. Kuan, H. L. Ong, and K. M. Ng, "Solving the feeder bus network design problem by genetic algorithms and ant colony optimization," *Advances in Engineering Software*, vol. 37, no. 6, pp. 351–359, 2006.

[5] C. D. Tarantilis, A. K. Anagnostopoulou, and P. P. Repoussis, "Adaptive path relinking for vehicle routing and scheduling problems with product returns," *Transportation Science*, vol. 47, no. 3, pp. 356–379, 2013.

[6] J. F. Campbell, "A continuous approximation model for time definite many-to-many transportation," *Transportation Research Part B: Methodological*, vol. 54, no. 7, pp. 100–112, 2013.

[7] K. Huang, G. H. d. A. Correia, and K. An, "Solving the station-based one-way carsharing network planning problem with relocations and non-linear demand," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 1–17, 2018.

[8] W. He, K. Hwang, and D. Li, "Intelligent carpool routing for urban ridesharing by mining GPS trajectories," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2286–2296, 2014.

[9] B. Eksioglu, A. V. Vural, and A. Reisman, "The vehicle routing problem: a taxonomic review," *Computers & Industrial Engineering*, vol. 57, no. 4, pp. 1472–1483, 2009.

[10] A. Verma and S. L. Dhingra, "Developing integrated schedules for urban rail and feeder bus operation," *Journal of Urban Planning and Development*, vol. 132, no. 3, pp. 138–146, 2006.

[11] A. Fügenschuh, "Solving a school bus scheduling problem with integer programming," *European Journal of Operational Research*, vol. 193, no. 3, pp. 867–884, 2009.

[12] M. E. Bruni, F. Guerriero, and P. Beraldi, "Designing robust routes for demand-responsive transport systems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 70, no. 1, pp. 1–16, 2014.

[13] T. Liu and A. Ceder, "Analysis of a new public-transport-service concept: customized bus in China," *Transport Policy*, vol. 39, pp. 63–76, 2015.

[14] Z. Liu, S. Wang, and Q. Meng, "Optimal joint distance and time toll for cordon-based congestion pricing," *Transportation Research Part B: Methodological*, vol. 69, no. 4, pp. 81–97, 2014.

[15] J.-F. Cordeau and G. Laporte, "The dial-a-ride problem: models and algorithms," *Annals of Operations Research*, vol. 153, no. 1, pp. 29–46, 2007.

[16] C. Dhaenens, J. Lemesre, and E. G. Talbi, "K-PPM: a new exact method to solve multi-objective combinatorial optimization problems," *European Journal of Operational Research*, vol. 200, no. 1, pp. 45–53, 2010.

[17] A. Przybylski, X. Gandibleux, and M. Ehrgott, "Two phase algorithms for the bi-objective assignment problem," *European Journal of Operational Research*, vol. 185, no. 2, pp. 509–533, 2008.

[18] M. Visée, J. Teghem, M. Pirlot, and E. L. Ulungu, "Two-phases method and branch and bound procedures to solve the bi–objective knapsack problem," *Journal of Global Optimization*, vol. 12, no. 2, pp. 139–155, 1998.

[19] D. Huang, Y. Gu, S. Wang, Z. Liu, and W. Zhang, "A two-phase optimization model for the demand-responsive customized bus network design," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 1–21, 2020.

[20] K. Huang, K. An, and G. H. de Almeida Correia, "Planning station capacity and fleet size of one-way electric carsharing systems with continuous state of charge functions," *European Journal of Operational Research*, vol. 287, no. 3, pp. 1075–1091, 2020.

[21] B. Boyacı, K. G. Zografos, and N. Geroliminis, "An integrated optimization-simulation framework for vehicle and personnel relocations of electric carsharing systems with reservations," *Transportation Research Part B: Methodological*, vol. 95, pp. 214–237, 2017.

[22] Z. Liu, S. Wang, W. Chen, and Y. Zheng, "Willingness to board: a novel concept for modeling queuing up passengers," *Transportation Research Part B: Methodological*, vol. 90, pp. 70–82, 2016.

[23] Z. Liu, S. Wang, K. Huang, J. Chen, and Y. Fu, "Practical taxi sharing schemes at large transport terminals," *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, pp. 596–616, 2019.

[24] D. Jorge, G. H. A. Correia, and C. Barnhart, "Comparing optimal relocation operations with simulated relocation policies in one-way carsharing systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 4, pp. 1667–1675, 2014.

[25] D. Jorge, G. Molnar, and G. H. de Almeida Correia, "Trip pricing of one-way station-based carsharing networks with zone and time of day price variations," *Transportation Research Part B: Methodological*, vol. 81, pp. 461–482, 2015.

[26] K. Sivakumaran, Y. Li, M. J. Cassidy, and S. Madanat, "Cost-saving properties of schedule coordination in a simple trunk-and-feeder transit system," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 1, pp. 131–139, 2012.

[27] Z. Xiang, C. Chu, and H. Chen, "A fast heuristic for solving a large-scale static dial-a-ride problem under complex constraints," *European Journal of Operational Research*, vol. 174, no. 2, pp. 1117–1139, 2006.

[28] Z. Xiang, C. Chu, and H. Chen, "The study of a dynamic dial-a-ride problem under time-dependent and stochastic environments," *European Journal of Operational Research*, vol. 185, no. 2, pp. 534–551, 2008.

WILEY | Hindawi

Research Article

Estimation of Disease Transmission in Multimodal Transportation Networks

Yu Zheng (ID)

School of Economic and Management, Tongji University, Shanghai 200092, China

Correspondence should be addressed to Yu Zheng; 1710270@tongji.edu.cn

Mathematical models are important methods in estimating epidemiological patterns of diseases and predicting the consequences of the spread of diseases. Investigation of risk factors of transportation modes and control of transportation exposures will help prevent disease transmission in the transportation system and protect people's health. In this paper, a multimodal traffic distribution model is established to estimate the spreading of virus. The analysis is based on the empirical evidence learned from the real transportation network which connects Wuhan with other cities. We consider five mainstream travel modes, namely, auto mode, high-speed railway mode, common railway mode, coach mode, and flight mode. Logit model of economics is used to predict the distribution of trips and the corresponding diseases. The effectiveness of the model is verified with big data of the distribution of COVID-19 virus. We also conduct model-based tests to analyze the role of lockdown on different travel modes. Furthermore, sensitivity analysis is implemented, the results of which assist in policy-making for containing infection transmission through traffic.

1. Introduction

Despite tremendous efforts to reduce and control infectious diseases, infections continue to be a global threat to worldwide public health. Understanding the virus propagation is quite essential for the implementation of antivirus methods. While research studies about the antivirus policy have been extensively investigated, the viewpoint from the perspective of the propagation along transportation modals is relatively ignored. Consideration of risk factors of transportation modes and control of transportation exposures will help prevent disease transmission in the transportation system and protect people's health. When an infectious disease case occurs at a location, investigators need to understand the mechanisms of disease propagation in the transportation network.

On December 31, 2019, the outbreak of novel coronavirus was first reported in China. The global outbreak of COVID-19 was mainly caused by transmission through different transportation modes. To prevent the spreading of virus, all the transportation system from Wuhan to the outside was closed in the morning of January 23, 2020. On January 30, 2020, the WHO (World Health Organization) declared a global emergency. On March 11, WHO declared the COVID-19 outbreak to be a global pandemic. For weeks after the first reports of a mysterious new virus of COVID-19, millions of people poured out of the central Chinese city, cramming onto buses, trains, and planes as the first wave of China's great Lunar New Year migration broke across the nation, and some of them are virus carriers. The travel patterns broadly track with the early spread of the virus. The majority of confirmed cases and deaths have occurred in China, within Hubei province, followed by high numbers of cases in central China, with pockets of infections in Chongqing, Shanghai, and Beijing as well. The initial spread of travelers to provinces in central China is with large pools of migrant workers. There might be a "high correlation" between the early spread of coronavirus cases and the distribution of travel destinations. The atmosphere in the transportation vessels is closed, and it is easy for the virus to spread. And the transmission speed is different in different traffic modals, due to the different air fluency in the traffic vessels.

Mathematical models have become important tools in epidemiology in understanding epidemiological patterns of diseases and predicting the consequences of the introduction of public health interventions to control the spread of diseases. There are two lines of studies in epidemics spreading. The first line is the spreading model of differential equation, and the second line is the complex network theory. In the literature, there are three spreading models widely used in modeling virus transmission, namely, SIR model, SIS model, and SI model (acronyms such as M, S, E, I, and R are often used for the epidemiological classes. The class M represents individuals with passive immunity. The class S represents susceptible individuals who can become infected. The class E represents the exposed individuals in the latent period, who are infected but not yet infectious. The class I represents the individuals of infective, who are infectious in the sense that they are capable of transmitting the infection. The class R represents recovered individuals with permanent infection-acquired immunity. The choice of which epidemiological class to include in a model depends on the characteristics of the particular disease being modeled and the purpose of the model) [1–4]. To solve the models, three kinds of algorithms have been developed based on percolation theory [5, 6], mean field theory [7, 8], and Markov chain theory [9, 10].

Researchers also developed models to investigate the propagation of different types of viruses including some nonbiological viruses, such as the computer virus, the flash disk virus, the Bluetooth phone virus, and the email virus. Otero-Muras et al. presented a systematic approach to the biochemical network dynamic analysis and control based on both thermodynamic and control theoretic tools [11]. Based on biological control strategy in pest management, Pang and Chen constructed a pest-epidemic model with impulsive control, i.e., periodically spraying microbial pesticide and releasing infected pests at different fixed moments [12]. Jin and Wang developed a new dynamic propagation model of FD-SEIR, namely, flash disk virus susceptible-exposed-infectious-recovered, which is embodied by introducing the FD state and new propagation rate [13]. Huang et al. developed an epidemic model of Bluetooth phone virus [14]. Li et al. formulated a novel deterministic SEIS model for the transmission of email viruses in growing communication networks [15]. Jackson and Chen-Charpentier presented two plant virus propagation models, one with no delays and the other with two delays [16]. Jia and Lv established a stochastic rumor propagation model. Sufficient conditions for extinction and persistence in the mean of the rumor have been examined [17]. Zhang et al. established a spreading model based on contact strength and SI model, and a weighted network with community structure based on a network model proposed by Barrat et al. [18].

In the following, a multimodal traffic distribution model is established to estimate the spreading of virus. The analysis is based on the empirical evidence learned from the real transportation network which connects Wuhan with other cities. Five travel modes are considered, namely, auto mode, high-speed railway mode, common railway mode, coach mode, and flight mode. Logit model of economics is used to predict the distribution of trips and the corresponding diseases. The effectiveness of the model is verified with big data of the distribution of COVID-19 virus. The main contributions of the paper are in four aspects. First, we propose a multimodal traffic distribution model using data of the real transportation system. Second, we study the relation between the state of disease transmission and the traffic flows distribution based on the numerical results of the proposed model and the big data of the distribution of COVID-19 virus. Third, we use the model to predict the role of lockdown on different transport means and analyze its impact on the disease transmission. Fourth, we present a sensitivity analysis for the proposed model and derive various transportation improvement policies to control large-scale transportation exposure.

The remainder of this paper is organized as follows. Section 2 establishes a multimodal traffic distribution model to estimate the spreading of virus. The proposed model is validated in Section 3, using a real traffic distribution from Wuhan to other regions in China during the outbreak of COVID-19. Conclusions are made in Section 4.

2. Multimodal User Equilibrium Model

2.1. Multimode Travel Cost Functions. The multimode travel cost functions are based on the empirical evidence learned from the real transportation network connecting Wuhan to other cities. We consider five mainstream travel modes, namely, auto mode, high-speed railway mode, common railway mode, coach mode, and flight mode (Tables 1 and 2). Let c_m^d represent the travel cost in the travel mode m ($m \in M$, $M = \{$auto, high–speed railway, common railway, coach, flight$\}$) to a destination region indexed by a region name d ($d \in D$, $D = \{$Xiaogan, Huanggang, Jingzhou, Xianning, E'zhou, Xiangyang, Huangshi, Jingmen, Suizhou, Xiantao, Yichang, Tianmen, Shiyan, Enshi, Qianjiang, Henan, Hunan, Anhui, Jiangxi, Guangdong, Jiangsu, Chongqing, Sichuan, Shandong, Zhejiang, Hebei, Fujian, Beijing, Guangxi, Shanxi, Shanghai, Shanxi, Guizhou, Yunnan, Hainan, Gansu, Liaoning, Heilongjiang, Xinjiang, Inner Mongolia, Jilin, Tianjin, Ningxia, Qinghai, Tibet, Hong Kong, Macao, Taiwan$\}$). The set of destination regions D includes 15 cities within the province of Hubei and 33 province-level regions in China. We further define the set of 15 cities within the province of Hubei as D_{in} ($D_{\text{in}} = \{$Xiaogan, Huanggang, Jingzhou, Xianning, E'zhou, Xiangyang, Huangshi, Jingmen, Suizhou, Xiantao, Yichang, Tianmen, Shiyan, Enshi, Qianjiang$\}$) and the set of other 33 province-level regions as D_{out} ($D_{\text{out}} = \{$Henan, Hunan, Anhui, Jiangxi, Guangdong, Jiangsu, Chongqing, Sichuan, Shandong, Zhejiang, Hebei, Fujian, Beijing, Guangxi, Shanxi, Shanghai, Shanxi, Guizhou, Yunnan, Hainan, Gansu, Liaoning, Heilongjiang, Xinjiang, Inner Mongolia, Jilin, Tianjin, Ningxia, Qinghai, Tibet, Hong Kong, Macao, Taiwan$\}$), and $D = D_{\text{in}} \cup D_{\text{out}}$.

The function of travel cost for each mode from Wuhan to a destination region indexed by d is described as follows:

TABLE 1: Automobile transportation parameters.

City	Province	Time	Distance	Toll	Gas fee
Xiaogan	Hubei	90	76.5	30	43
Huanggang	Hubei	77	75.3	30	42
Jingzhou	Hubei	180	220	90	123
Xianning	Hubei	90	92.6	30	52
E'zhou	Hubei	90	75	20	42
Xiangyang	Hubei	227	313	150	175
Huangshi	Hubei	100	100	40	56
Jingmen	Hubei	180	240	110	134
Suizhou	Hubei	120	170	80	95
Xiantao	Hubei	110	102	40	57
Yichang	Hubei	240	322	140	180
Tianmen	Hubei	120	142	50	80
Enshi	Hubei	420	519	250	291
Shiyan	Hubei	300	443	210	248
Qianjiang	Hubei	131	155	70	87
Shijiazhuang	Hebei	660	898	440	503
Taiyuan	Shanxi	720	944	450	529
Shenyang	Liaoning	1320	1812	890	1015
Changchun	Jilin	1560	2088	1030	1169
Harbin	Heilongjiang	1440	2354	1160	1318
Nanjing	Jiangsu	408	550	260	308
Hangzhou	Zhejiang	660	827	350	463
Hefei	Anhui	300	388	180	217
Fuzhou	Fujian	672	919	450	515
Nanchang	Jiangxi	330	355	170	199
Ji'nan	Shandong	600	864	420	484
Zhengzhou	Henan	360	514	250	288
Changsha	Hunan	289	345	150	193
Guangzhou	Guangdong	672	955	480	535
Haikou	Hainan	1152	1566	750	877
Chengdu	Sichuan	840	1130	550	633
Guiyang	Guizhou	720	1011	510	566
Kunming	Yunnan	1170	1558	760	872
Xi'an	Shanxi	510	740	360	414
Lanzhou	Gansu	930	1360	670	762
Xi'ning	Qinghai	1188	1594	790	893
Taibei	Taiwan	—	—	—	5600
Beijing	—	780	1174	580	657
Tianjin	—	720	1144	560	641
Shanghai	—	606	825	390	462
Chongqing	—	690	897	440	502
Hohhot	Inner Mongolia	960	1380	650	773
Nanning	Guangxi	810	1209	590	677
Lhasa	Tibet	3060	3482	1110	1950
Yinchuan	Ningxia	966	1448	710	811
Urumqi	Xinjiang	2160	3267	1600	1830
Hong Kong	—	840	1107	510	620
Macao	—	960	1210	600	678

Note that, in this part, the unit of measurement is kilometers for the distance, minutes for the time, and CNY for all kinds of tolls and fees.

(1) Auto mode:

$$c_{auto}^d = \left(\frac{vot_{auto} T_{auto}^d + \varepsilon_{auto} l_d + P_{highway}^d}{n} \right), \quad \forall d \in D, \quad (1a)$$

where ε_{auto} represents the cost of gasoline consumed per kilometer and T_{auto}^d denotes the auto travel time. The cost function c_{auto}^d consists of three terms. The first one stands for the monetary cost of travel time captured by the product of the value of in-vehicle travel time vot_{auto} and the travel time T_{auto}^d; the second one is the cost of gasoline consumed by this trip; the third one is the highway tolls charged along the highway captured by the product of the highway toll charged per kilometer and the total highway length. The average vehicle occupancy n is the average number of occupants in a vehicle. We set the n-piece of auto utility function taking account of the actual traffic situation. The Transport Bureau of

TABLE 2: Transportation parameters of high-speed railway, common railway, coach, and flight.

City	Province	High-speed railway		Common railway		Coach		Flight	
		Time	Ticket fee	Time	Ticket fee	Time	Ticket fee	Time	Ticket fee
Xiaogan	Hubei	30	58	60	14.5	60	32	—	—
Huanggang	Hubei	90	22	37	20	—	—	—	—
Jingzhou	Hubei	90	76	89	32.5	300	60	—	—
Xianning	Hubei	24	40	60	12.5	60	28	—	—
E'zhou	Hubei	24	20	80	12.5	60	28	—	—
Xiangyang	Hubei	90	130	190	50.5	240	88	—	—
Huangshi	Hubei	37	30	100	16.5	120	42	—	—
Jingmen	Hubei	—	—	190	40.5	180	101	—	—
Suizhou	Hubei	55	70	128	26.5	180	66	—	—
Xiantao	Hubei	60	50	—	—	90	35	—	—
Yichang	Hubei	150	121	310	53.5	270	105	—	—
Tianmen	Hubei	60	45	90	19.5	180	55	—	—
Enshi	Hubei	270	187	420	78	480	130	—	—
Shiyan	Hubei	144	217	330	72	390	135	—	—
Qianjiang	Hubei	90	65	—	—	150	57	—	—
Shijiazhuang	Hebei	240	415	540	124	—	—	450	1450
Taiyuan	Shanxi	390	486	1332	173.5	750	320	105	727
Shenyang	Liaoning	660	800	1320	217	1440	480	180	1820
Changchun	Jilin	680	904	1527	243	—	—	180	1800
Harbin	Heilongjiang	753	1012	1396	268.5	—	—	180	1800
Nanjing	Jiangsu	180	200	—	—	480	200	—	—
Hangzhou	Zhejiang	300	300	720	120	700	285	80	830
Hefei	Anhui	120	134	—	—	360	150	—	—
Fuzhou	Fujian	371	267	720	120	700	280	90	870
Nanchang	Jiangxi	150	120	344	53.5	390	120	—	—
Ji'nan	Shandong	360	525	720	130	780	280	95	1000
Zhengzhou	Henan	140	244	300	75	480	140	—	—
Changsha	Hunan	92	165	240	53.5	300	120	—	—
Guangzhou	Guangdong	260	464	750	138.5	960	340	110	1800
Haikou	Hainan	—	—	1440	250	1290	430	150	1700
Chengdu	Sichuan	560	375	990	185	960	330	120	1350
Guiyang	Guizhou	300	481	930	163.5	960	320	115	1000
Kunming	Yunnan	420	665	1373	217	1500	480	135	1660
Xi'an	Shanxi	270	455	900	135.5	560	240	85	1200
Lanzhou	Gansu	400	654	1200	190	1200	430	135	1330
Xi'ning	Qinghai	—	—	—	—	—	—	130	1300
Taibei	Taiwan	—	—	—	—	—	—	155	1400
Beijing	—	270	520	720	152.5	900	320	120	2200
Tianjin	—	300	525	840	156.5	960	300	115	1150
Shanghai	—	300	336	900	140	720	250	95	1880
Chongqing	—	390	279	540	140	780	280	95	1650
Hohhot	Inner Mongolia	—	—	1828	229	—	—	130	1050
Nanning	Guangxi	450	478	840	170	1050	320	120	1180
Lhasa	Tibet	—	—	—	—	—	—	230	1070
Yinchuan	Ningxia	—	—	1560	198	—	—	135	1300
Urumqi	Xinjiang	—	—	2310	345	—	—	260	2000
Hong Kong	—	280	679	—	—	—	—	130	1574
Macao	—	—	—	—	—	—	—	100	1380

Note that, in this part, the unit of measurement is kilometers for the distance, minutes for the time, and CNY for all kinds of tolls and fees.

Wuhan announced that the average vehicle occupancy of a privately owned automobile in Wuhan is 1.8 persons/vehicle. Besides, it was reported by the Spring Festival Transport Office of the province of Hubei that advantages such as the trip cost shared by several relatives and friends, larger space for luggage, and no need to transfer are attracting more and more individuals traveling back home by a private car during the Spring Festival travel season. It is therefore reasonable to assume that the number of occupants in a vehicle within the Spring Festival travel season should be no less than the average vehicle occupancy, i.e., 1.8 persons/vehicle. Taking account of this, the value of n is set as 2.

(2) High-speed railway mode:

$$c^d_{\text{high-speed rail}} = vot_{\text{high-speed rail}} T^d_{\text{high-speed rail}} + \tau^d_{\text{high-speed rail}}, \quad \forall d \in D, \tag{1b}$$

where $T^d_{\text{high-speed rail}}$ is the high-speed railway travel time and vot_{rail} is the value of time spent in the high-speed railway. The travel cost $c^d_{\text{high-speed rail}}$ consists of 2 terms: the first term is the cost of the travel time and the second term represents the high-speed railway ticket price.

(3) Common railway mode:

$$c^d_{\text{rail}} = vot_{\text{rail}} T^d_{\text{rail}} + \tau^d_{\text{rail}}, \quad \forall d \in D, \tag{1c}$$

where T^d_{rail} is the common railway travel time and vot_{rail} is the value of time spent in the common railway. The travel cost c^d_{rail} consists of 2 terms: the first term is the cost of the travel time and the second term represents the common railway ticket price.

(4) Coach mode:

$$c^d_{\text{coach}} = vot_{\text{coach}} T^d_{\text{coach}} + \tau^d_{\text{coach}}, \quad \forall d \in D, \tag{1d}$$

where T^d_{coach} is the coach travel time and vot_{coach} is the value of time spent in a coach. The travel cost c^d_{coach} consists of 2 terms: the first term is the cost of the travel time and the second term represents the coach ticket price.

(5) Flight mode:

$$c^d_{\text{flight}} = vot_{\text{flight}} T^d_{\text{flight}} + \tau^d_{\text{flight}}, \quad \forall d \in D, \tag{1e}$$

where T^d_{flight} is the travel time by taking a plane and vot_{bus} is the value of time spent in a flight. The travel cost c^d_{flight} consists of 2 terms: the first term is the cost of the travel time and the second term is the flight ticket price.

2.2. Multimodal User Equilibrium Model.

To cater for the consideration of both mode choice and destination choice, we propose a multimodal network user equilibrium model as follows:

$$\min \sum_{d \in D} \sum_{m \in M} \int_0^{q^d_m} \left(\frac{1}{\theta_d} \ln \frac{w}{q_d} + c^d_m(w) - \beta^d_m \right) dw$$
$$+ \sum_{d \in D} \int_0^{q_d} \left(\frac{1}{\delta} \ln \frac{w}{Q} - \alpha_d \right) dw, \tag{2a}$$

subject to

$$\sum_{d \in D} q_d = Q, \tag{2b}$$

$$\sum_{m \in M} q^d_m = q_d, \quad \forall d \in D, \tag{2c}$$

$$q_d \geq 0, \quad \forall d \in D, \tag{2d}$$

$$q^d_m \geq 0, \quad \forall d \in D, \forall m \in M, \tag{2e}$$

where θ_d is the impedance parameter associated with the travel mode choice to the destination region d, δ is the impedance parameter associated with destination choice, β^d_m is the exogenous attractiveness of the travel mode m to the destination d, and α_d is the exogenous attractiveness of the destination region d. q^d_m indicates travel demand in the travel mode m from Wuhan to the destination region d. q_d indicates travel demand from Wuhan to the destination region d. As to α_d, we develop the following weighted destination attractiveness measure to quantify the attractiveness of each destination to Wuhan:

$$\alpha_d = aH_d \cdot (L_d)^{k_d} + bP_d, \quad \forall d \in D, \tag{2f}$$

where H_d is the normalized historical demand distribution ratio of the destination region d; P_d is the normalized population of the destination region d; L_d is the normalized distance between the destination d and Wuhan; a is a weighting parameter for the parameter H_d; b is a weighting parameter for the parameter P_d; k_d is a weighting parameter for the parameter L_d. The weighting parameters a, b, and k_d reflect the extent of effects of historical demand distribution, population, and distance on the attractiveness of a destination region d. Learned from real-life experiences, the historical demand distribution ratio of a destination region d (H_d) is correlated oppositely to its travel distance to Wuhan (L_d). Furthermore, referring to equation (2l), for a destination region d, its generalized travel cost to Wuhan (u_d) varies incrementally with the travel time which is determined by the travel distance (L_d). To mitigate the interrelation between H_d and u_d, the term (aH_d) is multiplied by the term $(L_d)^{k_d}$.

We denote the original data of historical demand distribution ratio, population, and travel distance of a destination region d as h_d, p_d, and l_d. As h_d, p_d, and l_d are incommensurable, namely, all are measured in different units, they cannot be directly added and need to be normalized before the use of the weighted-sum method. To do so, we define H_d, P_d, and L_d as follows:

$$H_d = \frac{h_d}{\min(\mathbf{h_d})}, \quad \forall d \in D, \tag{2g}$$

$$P_d = \frac{p_d}{\min(\mathbf{p_d})}, \quad \forall d \in D, \tag{2h}$$

$$L_d = \frac{l_d}{\max(|\mathbf{l}|_\mathbf{d})}, \quad \forall d \in D, \tag{2i}$$

where $\min(\cdot)$ is a function to obtain the minimum item in a list, for example, $\min(|\mathbf{h}|_\mathbf{d})$ is to get the minimum item in the list $\mathbf{h_d}$. $\max(\cdot)$ is a function to obtain the maximum item in a list, for example, $\max(\mathbf{l_d})$ is to get the maximum item in the list $\mathbf{l_d}$. According to equation (2i), it is easy to get that $0 \leq L_d \leq 1$, from which we can infer that the value of the weighted distance parameter $(L_d)^{k_d}$ ranges from 0 to 1 for any k_d ($k_d > 0$). Besides, it can be learned from experiences

that passengers are more sensitive to travel cost in a short trip than in a long trip. This phenomenon has been studied in the area of stochastic traffic flow distribution [19]. In this work, passengers within the province of Hubei are more sensitive to the travel cost than those traveling out of the province of Hubei because of shorter travel distance. That is to say, the magnitude of travel cost takes a greater effect on the attractiveness of a destination region inside Hubei than that outside Hubei, meaning the value of $(L_d)^{k_d}$ should be greater for $d \in D_{in}$ than for $d \in D_{out}$. Along with the already known condition that $0 \le L_d \le 1$, it consequently requires the travel distance-related weighting parameter k_d taking a smaller value for $d \in D_{in}$ than $d \in D_{out}$. Furthermore, as the provincial capital of Hubei, Wuhan attracts a good many of migrant workers and students working or studying there each year for its abundant employment opportunities and diverse educational resources. The migrant population constitutes the majority of travel demands in the Spring Festival travel season in Wuhan. Without unexpected disruption, the historical traffic distribution of Wuhan in recent Spring Festival travel seasons will provide high-quality evidence for predicting the traffic distribution of this year. To reflect the significant impacts of the historical traffic distribution on the assessment of a destination's attractiveness, we suggest that the historical traffic distribution-related weighting parameter a takes a larger value than the population-related weighting parameter b. The detailed value setting for various parameters defined in this part can be found in Table 3.

The objective function (2a) is a two-level nested logit choice model to deal with the interrelated decisions in a multimodal network. The first level focuses on destination choice and the second level on mode choice. Equation (2b) ensures that the amount of flow assigned to different destination regions from Wuhan sums to the total travel demand Q which, in this work, amounts to 5,000,000. Equation (2c) represents the mode flow conservation constraint. Equations (2d) and (2e) are the nonnegativity conditions for destination demands and mode flows, respectively.

By deriving the first-order optimality conditions of the proposed program, we have the following nested logit model for destination choice and mode choice, respectively:

$$q_m^d = q_d \frac{\exp\left[-\theta_d\left(c_m^d - \beta_m^d\right)\right]}{\sum_{x \in M}\exp\left[-\theta_d\left(c_x^d - \beta_x^d\right)\right]}, \quad \forall d \in D, \forall m \in M, \tag{2j}$$

$$q_d = Q \frac{\exp\left[-\delta\left(u_d - \alpha_d\right)\right]}{\sum_{x \in D}\exp\left[-\delta\left(u_x - \alpha_x\right)\right]}, \quad \forall d \in D, \tag{2k}$$

where ud is users' perception of the generalized cost of traveling from the origin city Wuhan to the destination city d, which is computed as a "log-sum" of travel cost of each mode, i.e.,

$$u_d = -\frac{1}{\theta_d}\ln\sum_{m \in M}\exp\left[-\theta_d\left(c_m^d - \beta_m^d\right)\right], \quad \forall d \in D. \tag{2l}$$

To solve the nested logit model-based problem, one can first compute the generalized cost u_d ($\forall d \in D$) according to

equation (2l) and then carry out multiproportional traffic assignment (2j)-(2k) to obtain the combined destination distribution and modal split, i.e., q_m^d and q_d.

3. Case Study

The outbreak of COVID-19, which started in December last year, took Wuhan as the center and soon spread to all regions of China (including Hong Kong, Macao, and Taiwan). In the early morning of January 22, the province of Hubei launched level II emergency response to public health emergencies, and then cities in Hubei successively stopped public transportation. As of 11:00 on January 24, public transportation in 12 cities in Hubei had been shut down, including Wuhan, E'zhou, Xiantao, Zhijiang, Qianjiang, Huanggang, Chibi, Jingmen, Xianning, Huangshi, Dangyang, and Enshi, among which Wuhan, as the transport hub of more than 10 million people, temporarily closed its airports, rail stations, and all main roads out of town, as well as suspended public buses and subways. The government announced that citizens should not leave Wuhan without special reasons, and the lift of the lockdown will be announced separately. On January 26, the Information Office of the People's Government of Hubei held a press conference, pointing out that from the beginning of the Spring Festival to the closure of Wuhan, more than 5 million people left Wuhan, and more than 9 million remained in the city.

In this section, we will use the transportation model proposed in Section 2 to analyze the traffic flow distribution for the 5 million people outbound from Wuhan and then estimate the epidemic situation based on the demand distribution results. We are mainly concerned about the distribution of people within the province of Hubei as well as outside the province of Hubei. Figure 1 shows the map of the province of Hubei and 35 other regions of China, and Figure 2 shows the map of Wuhan and 16 other cities in the province of Hubei.

To facilitate the computation of the travel utility to a destination province outside Hubei, instead of calculating the travel utility to each city in the destination province, we only calculate the travel utility to the provincial capital city. For the calculation of the normalized historical demand distribution ratio parameter H_d in equation (2f), we collect the data of migration from Wuhan to other destination regions of the year 2017 on the Tencent social network's Spring Festival geographic positioning data platform. The data show that except several provinces including Henan, Hunan, Anhui, Jiangsu, and Guangdong, for other provinces, the majority of the traffic out of Wuhan flowed into their provincial capitals. As a result, we replace the population of a province by the population of its provincial capital for the computation of the parameter P_d in equation (2f). As to other five provinces, i.e., Henan, Hunan, Anhui, Jiangsu, and Guangdong, we use the sum of population of cities which occupied the most amount of immigration from Wuhan in 2017 instead of the population of the province. Besides, to obtain the travel distance parameter L_d in equation (2f), we use the road length from Wuhan to other destination regions to measure the travel distance. The data

TABLE 3: Parameters for nested logit model-based traffic assignment.

vot_{auto}	$vot_{high-speed\,rail}$	vot_{rail}	vot_{coach}	vot_{flight}
1.2	1.1	1.25	1.3	1.0
ε_{auto}	n	$\theta_d\,(\forall d \in D_{in})$	$\theta_d\,(\forall d \in D_{out})$	$\beta_m^d\,(\forall d \in D, \forall m \in M)$
0.6	2	0.05	0.01	0.0
δ	a	b	$k_d\,(\forall d \in D_{in})$	$k_d\,(\forall d \in D_{out})$
0.003	10	5	0.3	0.5

FIGURE 1: Wuhan and other 35 regions outside the province of Hubei.

FIGURE 2: Wuhan and other 16 cities in the province of Hubei.

of historical demand distribution, the population of cities and provinces, and the road distance from Wuhan to other destination regions can be found in Table 4. parameters for the computation of the nested logit model-based traffic assignment are listed in Table 3. Learned from the real traveling experiences, parameters are set as follows:

TABLE 4: The historical demand distribution data, the population of cities and provinces, and the road distance from Wuhan to other destination regions.

Destination	Historical demand distribution ratio	Population in millions	Road distance
Xiaogan	0.118615	4.8780	76.5
Huanggang	0.110331	6.2910	75.3
Jingzhou	0.055153	5.7442	220
Xianning	0.043188	2.4626	92.6
E'zhou	0.031015	1.0487	75
Xiangyang	0.036824	5.6140	313
Huangshi	0.033994	2.4293	100
Jingmen	0.029304	2.8737	240
Suizhou	0.028677	2.1622	170
Xiantao	0.029431	1.1660	102
Yichang	0.024785	4.1150	322
Tianmen	0.020131	1.4189	142
Enshi	0.016231	3.2903	519
Shiyan	0.017129	3.3830	443
Qianjiang	0.011413	0.9463	155
Hebei	0.01537	11.0312	898
Shanxi	0.00676	4.4619	944
Liaoning	0.00442	8.3160	1812
Jilin	0.00177	7.6770	2088
Heilongjiang	0.00246	10.8580	2354
Jiangsu	0.01780	25.3086	550
Zhejiang	0.01188	10.3600	827
Anhui	0.02734	24.5670	388
Fujian	0.00958	7.8000	919
Jiangxi	0.02451	5.5455	355
Shandong	0.01502	7.4604	864
Henan	0.07515	40.5440	514
Hunan	0.04200	32.6188	345
Guangdong	0.02316	45.9177	955
Hainan	0.00337	2.3023	1566
Sichuan	0.02397	16.3300	1130
Guizhou	0.00746	4.8819	1011
Yunnan	0.00627	6.8500	1558
Shanxi	0.01332	10.0037	740
Gansu	0.00433	3.7536	1360
Qinghai	0.00097	2.3871	1594
Beijing	0.01147	21.536	1174
Tianjin	0.00222	15.6183	1144
Shanghai	0.00792	24.2814	825
Chongqing	0.01638	31.2432	897
Inner Mongolia	0.00195	3.1260	1380
Guangxi	0.00840	7.5687	1209
Ningxia	0.00074	2.2931	3267
Xinjiang	0.00182	3.5058	1107

Note that, in this part, the unit of measurement is kilometers for the distance, minutes for the time, and CNY for all kinds of tolls and fees.

$0 < vot_{flight} < vot_{high-speed\,rail} < vot_{auto} < vot_{rail} < vot_{coach}$. Furthermore, in this study, we consider different travel cost sensitivities of passengers with different scale of travel path sizes. The related research results [19] reveal that passengers in short trip are more sensitive to travel distance or travel cost than those in long trip, which causes the value of θ_m^d for $\forall d \in D_{in}$ is 5 times that for $\forall d \in D_{out}$.

3.1. Demand Assignment. Based on the model proposed in Section 2, we calculate the traffic flows from Wuhan to other 48 destination regions which include 15 cities within Hubei, and 33 destination regions outside Hubei. Note that

we exclude several regions which include the Shennongjia Forest District in Hubei, the Diaoyu Islands, and the South China Sea Islands from the calculation of destination demand distribution for that the traffic flows of these regions are very small. Besides, the demand distribution in real condition is collected from the Baidu Migration Big Data Platform. The error ratio of estimation is defined as the ratio of the estimation error to the result in real condition. Data of the estimated demand distribution, the demand distribution in real condition, and the error of estimation are listed in Tables 5 and 6 for cities within the province of Hubei and destination regions outside Hubei, respectively.

TABLE 5: Demand distribution estimation and the error of estimation for cities inside the province of Hubei.

Destination	Real condition	Est. results	Error	Error ratio (%)
Xiaogan	690000.00	665702.75	−24297.25	−3.52
Huanggang	652000.00	612343.63	−39656.37	−6.08
Jingzhou	327000.00	320223.65	−6776.35	−2.07
Xianning	250500.00	274428.51	23928.51	9.55
E'zhou	198500.00	235207.15	36707.15	18.49
Xiangyang	196500.00	193181.29	−3318.71	−1.69
Huangshi	188500.00	236658.71	48158.71	25.55
Jingmen	165000.00	156216.92	−8783.08	−5.32
Suizhou	160500.00	199004.52	38504.52	23.99
Xiantao	148500.00	202067.37	53567.37	36.07
Yichang	140500.00	131028.35	−9471.65	−6.74
Tianmen	104000.00	183119.51	79119.51	76.08
Shiyan	93000.00	89343.10	−3656.90	−3.93
Enshi	90500.00	64965.07	−25534.93	−28.22
Qianjiang	57000.00	140509.64	83509.64	146.51

TABLE 6: Demand distribution estimation and the error of estimation for other province-level destination regions outside the province of Hubei.

Destination	Real condition	Est. results	Error	Error ratio (%)
Henan	284000.00	250214.37	−33785.63	−11.90
Hunan	174000.00	187131.66	13131.66	7.55
Anhui	113500.00	144540.75	31040.75	27.35
Jiangxi	106000.00	128569.12	22569.12	21.29
Guangdong	97000.00	44218.78	−52781.22	−54.41
Jiangsu	73000.00	88634.88	15634.88	21.42
Chongqing	63500.00	42893.78	−20606.22	−32.45
Sichuan	62000.00	20641.30	−41358.70	−66.71
Shandong	55000.00	27805.30	−27194.70	−49.44
Zhejiang	53500.00	45031.19	−8468.81	−15.83
Hebei	46500.00	40775.05	−5724.95	−12.31
Fujian	45500.00	36048.21	−9451.79	−20.77
Beijing	43000.00	26050.15	−16949.85	−39.42
Guangxi	39500.00	14846.45	−24653.55	−62.41
Shanxi	36000.00	37590.59	1590.59	4.42
Shanghai	33000.00	36512.91	3512.91	10.65
Shanxi	29500.00	22472.40	−7027.60	−23.82
Guizhou	27500.00	21864.32	−5635.68	−20.49
Yunnan	26500.00	8502.62	−17997.38	−67.91
Hainan	19000.00	2387.83	−16612.17	−87.43
Gansu	17500.00	8962.46	−8537.54	−48.79
Liaoning	16500.00	2702.27	−13797.73	−83.62
Heilongjiang	14000.00	1237.75	−12762.25	−91.16
Xinjiang	10000.00	68.07	−9931.93	−99.32
Inner Mongolia	9000.00	6731.88	−2268.12	−25.20
Jilin	8500.00	1706.55	−6793.45	−79.92
Tianjin	7500.00	18275.56	10775.56	143.67
Ningxia	4000.00	3969.71	−30.29	−0.76
Qinghai	3000.00	3140.77	140.77	4.69
Tibet	1000.00	4156.12	3156.12	315.61
Others	19500.00	18317.04	−1182.96	−6.07

Table 7 provides the results of the aggregated demand distribution ratio, which shows that, in both the real and estimation conditions, the traffic flows within the province of Hubei account for most part (about 70%) of the total demands. It is also shown in Table 7 that the aggregated error ratio of demand estimation for cities inside Hubei (6.99%) is smaller than that for destinations outside Hubei (15.73%), meaning it performs better in demand estimation within the province of

Hubei, and the aggregated error ratio of demand estimation for all destination regions is 18.60%. The demand distribution results which take a decreasing order are shown in Figure 3. The destination name marked with an asterisk denotes a city within the province of Hubei. Note that in Figure 3, we put the demand distribution results of Hong Kong, Macao, and Taiwan into one item named "others" for brevity. According to results in Figure 3, in both the real and estimation conditions,

TABLE 7: Results of the aggregated demand distribution ratio and the aggregated error ratio of the estimated demand distribution for all destinations, destinations inside Hubei, and destinations outside Hubei.

Destination range	Aggregated traffic distribution ratio		Aggregated error ratio of estimation (%)
	Real condition (%)	Estimation (%)	
All	—	—	18.60
Des. inside Hubei	69.24	74.08	6.99
Des. outside Hubei	30.76	25.92	15.73
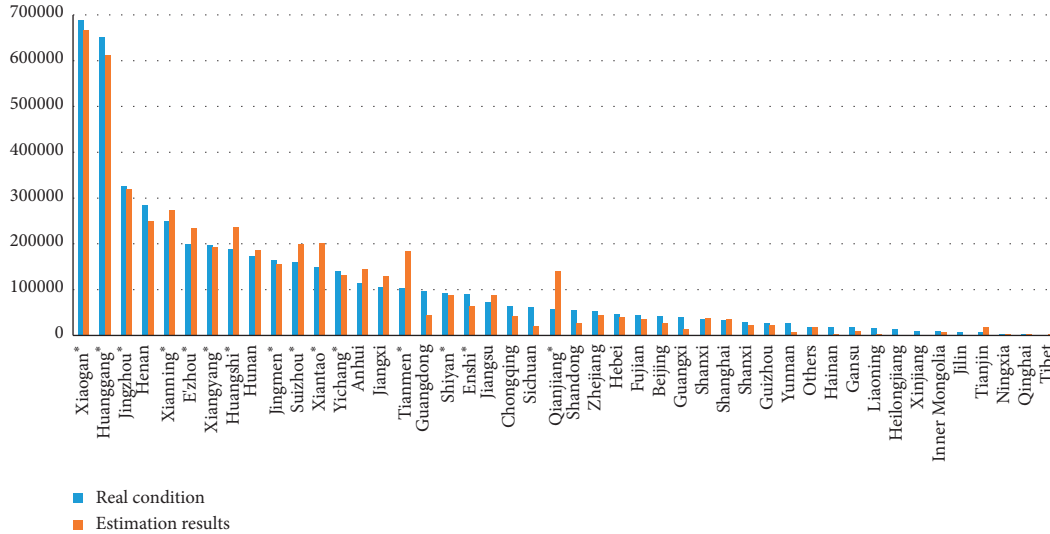
FIGURE 3: Traffic flow distribution results for the real condition.

Xiaogan, Huanggang, Jingzhou, Xianning, and E'zhou are the top five cities within the province of Hubei with the largest distribution of people, and Henan, Hunan, Anhui, and Jiangxi are the top four destination regions outside Hubei with the largest distribution of people. All results indicate that in a tolerable error range, our model delivers a desirable performance on the estimation of traffic flow distribution.

3.2. Numerical Estimation of Incidence Cases.
According to statistics released by the Chinese health authority, after March 18, all the increased confirmed cases in China are imported from overseas. Therefore, we use statistics of the day, March 18, to obtain the number of confirmed cases resulted from the travelers from Wuhan. The average incidence rate (γ) of people leaving Wuhan is defined as follows:

$$\gamma = \frac{f}{5,000,000}, \tag{3a}$$

where f is the number of confirmed cases nationwide apart from Wuhan. With $f = 31296$, we have $\gamma = 0.6259\%$. We then further estimate the number of incidence cases in different destination regions which is equal to $q_d * \gamma, \forall d \in D$. Results of the real number of incidence cases, the estimated number of incidence cases, the error of estimation for the number of incidence cases as well as the error ratio estimation for the number of incidence cases are listed in Tables 8 and 9 for destinations within and outside Hubei, respectively.

TABLE 8: Estimation of the number of incidence cases and the error of estimation for cities inside the province of Hubei.

Destination	Real condition	Est. results	Error	Error ratio (%)
Xiaogan	3518.00	4166.77	648.77	18.44
Huanggang	2907.00	3832.78	925.78	31.85
Jingzhou	1580.00	2004.34	424.34	26.86
Xianning	836.00	1717.70	881.70	105.47
E'zhou	1394.00	1472.21	78.21	5.61
Xiangyang	1175.00	1209.16	34.16	2.91
Huangshi	1015.00	1481.29	466.29	45.94
Jingmen	928.00	977.79	49.79	5.37
Suizhou	1307.00	1245.61	−61.39	−4.70
Xiantao	575.00	1264.78	689.78	119.96
Yichang	931.00	820.13	−110.87	−11.91
Tianmen	496.00	1146.18	650.18	131.09
Shiyan	672.00	559.22	−112.78	−16.78
Enshi	252.00	406.63	154.63	61.36
Qianjiang	198.00	879.48	681.48	344.18

Figure 4 illustratively presents results in Tables 9 and 10. It can be observed from Figure 4 that our estimation overestimates the number of incidence cases in most cities within Hubei, as well as two provinces, i.e., Henan and Hunan. According to results in Section 3.1, these cities/provinces are the destination regions with the largest traffic flow distributions. The fact of the lower incidence rate of these destination regions with the most immigration from Wuhan than the average incidence rate implies that

TABLE 9: Estimation of the number of incidence cases and the error of estimation for other province-level destination regions outside the province of Hubei.

Destination	Real condition	Est. results	Error	Error ratio (%)
Henan	1274.00	1566.14	292.14	22.93
Hunan	1018.00	1171.29	153.29	15.06
Anhui	990.00	904.71	−85.29	−8.62
Jiangxi	936.00	804.74	−131.26	−14.02
Guangdong	1415.00	276.77	−1138.23	−80.44
Jiangsu	633.00	554.78	−78.22	−12.36
Chongqing	577.00	268.48	−308.52	−53.47
Sichuan	543.00	129.20	−413.80	−76.21
Shandong	768.00	174.04	−593.96	−77.34
Zhejiang	1238.00	281.86	−956.14	−77.23
Hebei	319.00	255.22	−63.78	−19.99
Fujian	313.00	225.63	−87.37	−27.91
Beijing	537.00	163.05	−373.95	−69.64
Guangxi	254.00	92.93	−161.07	−63.41
Shanxi	248.00	235.29	−12.71	−5.13
Shanghai	404.00	228.54	−175.46	−43.43
Shanxi	133.00	140.66	7.66	5.76
Guizhou	146.00	136.85	−9.15	−6.26
Yunnan	176.00	53.22	−122.78	−69.76
Hainan	168.00	14.95	−153.05	−91.10
Gansu	136.00	56.10	−79.90	−58.75
Liaoning	127.00	16.91	−110.09	−86.68
Heilongjiang	484.00	7.75	−476.25	−98.40
Xinjiang	76.00	0.43	−75.57	−99.44
Inner Mongolia	75.00	42.14	−32.86	−43.82
Jilin	93.00	10.68	−82.32	−88.51
Tianjin	141.00	114.39	−26.61	−18.87
Ningxia	75.00	24.85	−50.15	−66.87
Qinghai	18.00	19.66	1.66	9.21
Tibet	1.00	26.01	25.01	2501.40
Hong Kong	233.00	114.65	−118.35	−50.79
Macao	155.00	70.40	−84.60	−54.58
Taiwan	11.00	31.83	20.83	189.38
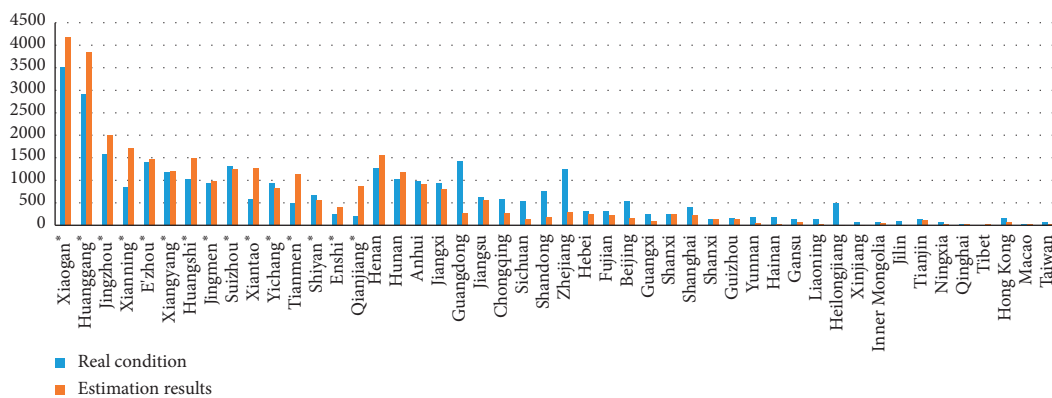
FIGURE 4: Comparison of the estimated number and the real number of incidence cases.

measures adopted by the Chinese government played an effective role in preventing a more serious situation from developing. Measures in the prevention and control of the epidemic involve lockdown on public transportations in 12 cities in Hubei, a two-week mandatory self-quarantine for people immigrated from Wuhan, residential community-based management, constructing temporary treatment centers, the centralized schedule of medical staff and supplies to the scarce areas, and so on. At the same time, results in Figure 4 also reveal that the model-based results underestimate the incidence rate of several destination regions including Guangdong, Chongqing, Sichuan, Shandong, Zhejiang, Beijing, Shanghai, Heilongjiang, and Hong Kong. This is because high frequent commercial activities that

TABLE 10: Mode flow distribution results of cities inside the province of Hubei.

City	Road	High-speed rail	Rail	Coach	Flight
Xiaogan	170285.59	178569.34	223626.44	93221.37	—
Huanggang	161473.63	25124.46	425745.54	—	—
Jingzhou	5018.37	45290.82	269914.25	0.22	—
Xianning	33196.64	151630.82	61341.02	28260.04	—
E'zhou	25378.08	188460.03	8447.63	12921.41	—
Xiangyang	3884.20	174498.63	14686.33	112.13	—
Huangshi	28096.44	200465.62	7468.12	628.54	—
Jingmen	129761.11	—	24305.32	2150.49	—
Suizhou	30586.40	155330.54	13007.88	79.70	—
Xiantao	87947.54	90625.94	—	23493.89	—
Yichang	29596.79	101257.41	94.67	79.47	—
Tianmen	26117.57	109137.47	47827.86	36.61	—
Shiyan	9449.63	79120.77	771.80	0.90	—
Enshi	5196.79	59325.61	441.77	0.90	—
Qianjiang	76189.15	62691.08	—	1629.41	—

involve face-to-face or close contact with other people lead to the higher incidence rates of the economically developed provinces, such as Guangdong, Zhejiang, Beijing, Shanghai, and Hong Kong. It is also interesting to see that the province of Heilongjiang, far away from Wuhan, not as commercially active as the provinces mentioned above, is featured by its high incidence rate. The high incidence rate of Heilongjiang may be attributed to the mass contact transmission of virus in gathering activities. According to news reported in Heilongjiang, to the date of February 7, there had been 48 family aggregating activities which were the source of 194 cases of cluster infection.

In Table 11, we compute the aggregated error ratio of estimation for the number of incidence cases which is equal to 39.52%, almost two times of the aggregated error ratio of estimation for the traffic flow distribution (18.60%), indicating that the spread of the epidemic is not linear with respect to the model-based traffic flow distribution. For cities within Hubei, the aggregated number of incidence case distribution ratio in real condition (56.40%) is much lower than the estimation (74.08%) while for provinces outside Hubei, the aggregated number of incidence case distribution ratio in real condition (43.60%) is much higher than the estimation (25.92%). This is because the spread of disease within Hubei is well controlled by means of transport restriction, medical assistance, and other effective methods while the high economic activity frequency as well as the high occurrence of mass gatherings in some provinces outside Hubei will potentially increase the incidence rate outside Hubei.

3.3. Mode Flow Distribution. Public transport as the main mode of transportation in big cities carries the highest risk of transmission of infection for a number of reasons. The high density of passengers confined in relatively small spaces was the primary cause. Besides, the in-vehicle air conditioning system featured by the low ventilation rates makes it easy for virus to spread. And the indirect infection from the contaminated public facilities in transport vessels is also one of the major danger sources. Furthermore, for passengers

taking a long trip, multiple public transportation transfers are often involved, the fact of which potentially increases the incidence rate. In contrast, self-driving or taking a ride in a privately owned vehicle has several advantages over public transport in containing the transmission of infection. First, passengers are separated by vehicles. The spatial isolation reduces the risk of cross infection. Second, in the self-driving travel mode, passengers drive to destinations directly without any transfer most of the time. Third, people who are friends or familiar with each other often travel together in a privately owned vehicle. It is easy for them to learn the health condition of each other which helps to raise their awareness of health security and as a result mitigates the risk of infection. Comparisons of different transportation means' impacts on the virus spreading reveal that it is important to enhance the epidemic prevention from the perspective of public transport control.

In this section, we first calculate the mode flow distribution based on the proposed model. The mode flow distributions of each destination region are listed in Tables 10 and 12. The results of the aggregated mode flow ratio for destinations inside Hubei as well as outside Hubei are shown in Figure 5. It can be seen from Figure 5 that, for destinations both inside and outside Hubei, public transports are the mainstream transportation means accounting for about 80% of the total demands. Besides, the most popular travel mode of public transport is the high-speed railway for trips both inside and outside Hubei, which indicates that enhanced measurements, such as disinfection and disease detection, should be adopted by the high-speed railway transportation system. Furthermore, the proportion of aggregated mode flow ratio of the common railway inside Hubei (29.63%) is much higher than that outside Hubei (13.55%), indicating that, for trips from Wuhan to cities inside Hubei, extra efforts should also be paid on the epidemic control in the common railway transportation system.

To contain the COVID-19 outbreak, many countries have implemented flight restrictions to China. At the same time, China itself has imposed a lockdown of the transportation system of Wuhan as well as the entire Hubei province. In this context, it is reasonable to investigate how

TABLE 11: Results of the aggregated number of incidence case distribution ratio and the aggregated error ratio of the number of estimated incidence case distribution for all destinations, destinations inside Hubei, and destinations outside Hubei.

Destination range	Aggregated number of incidence case distribution ratio		Aggregated error ratio of the estimation (%)
	Real condition (%)	Estimation (%)	
All	—	—	39.52
Within Hubei	56.40	74.08	30.36
Outside Hubei	43.60	25.92	39.94

TABLE 12: Mode flow distribution results of other province-level destination regions outside the province of Hubei.

Destination	Road	High-speed rail	Rail	Coach	Flight
Henan	44850.03	118548.74	81885.68	4929.93	—
Hunan	39450.61	92948.86	43774.41	10957.78	—
Anhui	32300.71	107671.97	—	4568.06	—
Jiangxi	28926.26	82248.90	13420.05	3973.91	—
Guangdong	5998.14	36177.92	2020.71	21.68	0.33
Jiangsu	16969.83	70065.96	—	1599.09	—
Chongqing	3815.81	26847.00	12063.14	166.99	0.84
Sichuan	4396.64	15655.80	343.30	115.42	130.14
Shandong	16523.83	6598.24	3179.75	345.35	1158.13
Zhejiang	5446.38	36087.67	1047.02	255.62	2194.49
Hebei	3402.03	26799.17	10573.71	—	0.13
Fujian	3006.54	29630.07	1349.57	346.38	1715.65
Beijing	1164.38	21849.67	2971.85	64.20	0.05
Guangxi	1994.34	10993.61	1415.26	25.41	417.83
Shanxi	21077.62	13574.76	131.76	2740.70	65.76
Shanghai	7129.73	28941.40	113.64	328.00	0.13
Shanxi	4241.40	5456.97	1.04	258.44	12514.56
Guizhou	2971.35	17850.25	164.81	24.04	853.86
Yunnan	122.45	8364.34	5.24	0.08	10.50
Hainan	2237.15	—	26.92	26.92	96.83
Gansu	978.49	7754.39	36.45	3.31	189.81
Liaoning	183.87	2345.41	149.94	2.56	20.50
Heilongjiang	33.87	751.06	267.06	—	185.77
Xinjiang	2.69	—	1.79	—	63.58
Inner Mongolia	1385.89	—	0.02	—	5345.97
Jilin	22.74	1600.39	23.20	—	60.22
Tianjin	2057.90	15236.40	689.83	38.92	252.51
Ningxia	2615.52	—	2.36	—	1351.83
Qinghai	545.38	—	—	—	2595.38
Tibet	—	—	—	—	4156.12
Hong Kong	2933.43	8307.63	—	—	6.39
Macao	4666.53	—	—	—	419.13
Taiwan	—	—	—	—	1983.93

the mode flow distribution changes with different outbound transport restrictions in Wuhan. We will use the proposed nested logit model to analyze the role of lockdown on each transport means in the following content.

3.3.1. Lockdown Test. Table 13 shows the results of the aggregated demands ratio inside Hubei under cases applying lockdown on different travel modes and it reveals that a lockdown on any travel mode will lead to an increase of the aggregated demands ratio inside Hubei, among which shutting down the high-speed railway will cause the maximum rise of the aggregated demands ratio inside Hubei

from 73.87% to 78.15%. This indicates that a lockdown on any travel mode will not make a big difference to the change of the aggregated demand distribution between destinations inside and outside Hubei. We then check the effects of transport restriction on the change of mode flow distribution, and related results are listed in Tables 14 and 15 for destinations inside Hubei and outside Hubei, respectively. From Tables 14 and 15, it can be seen that, for destinations both inside and outside Hubei, lockdown on a certain transportation means leads to the growth of traffic flows of other travel modes, and particularly lockdown on the high-speed railway has the most prominent impact on the traffic flow increment of other travel modes, indicating that in the

(a)

(b)
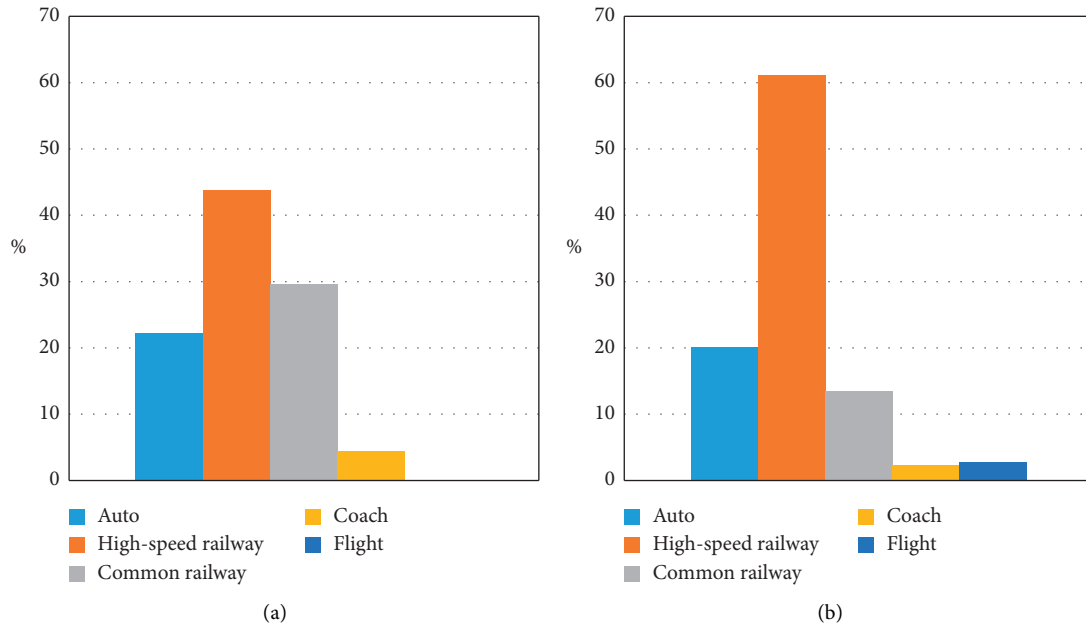
Figure 5: Aggregated mode flow distribution ratio of destination regions: (a) inside Hubei and (b) outside Hubei.

Table 13: Results of the aggregated traffic flow ratio of cities inside Hubei under cases applying lockdown on different travel modes.

Original case (%)	Locked down travel mode				
	Automobile (%)	High-speed railway (%)	Common railway (%)	Coach (%)	Flight (%)
74.08	75.06	78.63	74.41	74.16	74.34

Table 14: Results of the aggregated mode flow increment in percentage under cases applying lockdown on different travel modes for destinations inside Hubei.

Aggregated mode flow increment (%)	Locked down travel mode			
	Automobile (%)	High-speed railway (%)	Common railway (%)	Coach (%)
Automobile	—	121.59	63.65	5.49
High-speed railway	25.54	—	32.78	4.31
Common railway	35.29	62.22	—	4.02
Coach	42.98	102.42	36.47	—

Table 15: Results of the aggregated mode flow increment in percentage under cases applying lockdown on different travel modes for destinations outside Hubei.

Aggregated mode flow increment (%)	Locked down travel mode				
	Automobile (%)	High-speed railway (%)	Common railway (%)	Coach (%)	Flight (%)
Automobile	—	127.41	14.82	2.42	4.13
High-speed railway	19.98	—	14.34	1.94	1.37
Common railway	19.77	98.77	—	2.32	0.50
Coach	28.28	114.78	17.71	—	1.13
Flight	29.38	63.29	4.03	0.77	—

case of a lockdown on the high-speed railway, enforcement on the control of transportation exposures should be conducted for all the other public transport systems. It also reveals that a lockdown on a certain travel mode may cause different extent of aggregated mode flow increment of other travel modes. For example, the common railway restriction has the most significant impact on the increase of the

aggregated mode flows of automobile (63.65%) for destinations inside Hubei. And the automobile restriction leads to higher aggregated mode flow growth of coach (42.98%) for destinations inside Hubei than any other aggregated mode flow increment. This indicates that it is important to measure the magnitude of correlation between lockdown on a certain travel mode and the traffic flow increase of other

TABLE 16: Derivatives of the aggregated mode flows of auto as well as the aggregated demands of destination regions with high incidence rates with respect to different input parameters.

Aggr. mode flow	$\partial(\cdot)/\partial P_{\text{highway}}$	$\partial(\cdot)/\partial\tau_{\text{high-speed rail}}$	$\partial(\cdot)/\partial\tau_{\text{rail}}$	$\partial(\cdot)/\partial\tau_{\text{coach}}$	$\partial(\cdot)/\partial\tau_{\text{flight}}$	$0.1 * \partial(\cdot)/\partial n$
Q_{auto}	-14934.37	16010.48	11616.38	2172.01	69.88	26895.38
$Q_{\text{high incidence}}$	21.96	-232.22	168.19	21.8	3.76	1954.73

travel modes, and the public transport mode which has a high correlation with the lockdown needs intensified management to contain virus spreading through transportation.

3.4. Sensitivity Analysis. In this work, the logit-based probability expression for both destination and mode choice ensures that the solution to the lower-level programming is unique. Hence, the standard sensitivity analysis method for nonlinear programming problem can be used directly to derive the sensitivity information. The detailed derivation can be referred to Yang and Chen [20] and Yang et al. [21]. In this section, we conduct sensitivity analyses to explore how changes in input parameters including road tolls, high-speed railway ticket fees, common railway ticket fees, coach ticket fees, flight ticket fees, and the average vehicle occupancy affect certain traffic flows we are interested in. To ease the work of analysis, we investigate the change of traffic flows with respect to the same amount of perturbations of a particular parameter for all destinations rather than for each destination, respectively. For example, the term of derivative $\partial(Q_{\text{auto}})/\partial P_{\text{highway}}$ indicates the change of the traffic flow Q_{auto} with respect to an increase of 1 Chinese Yuan (CNY) in the road tolls for all destinations, different from the derivative term $\partial(Q_{\text{auto}})/\partial P_{\text{highway}}^{d}$ ($\forall d \in D$) which represents the change of the traffic flow Q_{auto} with respect to an increase of 1 CNY in the road toll for the specific destination indexed by d.

As aforementioned, compared with other public transport modes, traveling in privately owned vehicles contributes to less transportation exposures. As a result, measures taken to encourage traffic flows shifting from the public transport modes to the auto mode will mitigate transmission risks. We check the derivatives of the aggregated mode flows of auto which is defined as Q_{auto} with respect to perturbations of input parameters in Table 16. According to the results in Table 16, an increase of n has the most direct positive impact on Q_{auto}, which is followed by an increase of $\tau_{\text{high-speed rail}}$. This indicates that the increase of the average vehicle occupancy n and the increase of high-speed railway ticket fee $\tau_{\text{high-speed rail}}$ for all destinations are the most effective way to boost the aggregated mode flow of auto:

$$Q_{\text{high incidence}} = q_{\text{Zhejiang}} + q_{\text{Jiangsu}} + q_{\text{Guangdong}} + q_{\text{Beijing}}$$
$$+ q_{\text{Shanghai}} + q_{\text{Hong Kong}} + q_{\text{Heilongjiang}}.$$

(3b)

As we discussed in Section 3.2, the actual incidence rates in economically developed destinations as well as destinations with high occurrence of big gathering activities are

much higher than the estimated incidence rates. It is a natural thought to prevent the virus spreading from seriously developing by curbing the demands of these destinations. The aggregated demands of the economically developed destinations as well as destinations with high occurrence of big gathering activities is defined as $Q_{\text{high incidence}}$ in equation (3b), which is the sum of demands of multiple destinations involving Zhejiang, Jiangsu, Guangdong, Beijing, Shanghai, Hong Kong, and Heilongjiang. We check the sensitivity of the aggregated demands of these destinations ($Q_{\text{high incidence}}$) with respect to different parameters in Table 16. One useful application of the derivatives of $Q_{\text{high incidence}}$ is to identify effective measures to induce negative growth of $Q_{\text{high incidence}}$. The negative value of $\partial Q_{\text{high incidence}}/\partial(\cdot)$ indicates that an increase in a particular parameter leads to a decline in $Q_{\text{high incidence}}$. In this study, decreasing n will cause the most decline in $Q_{\text{high incidence}}$, and the second largest decline is generated by increasing $\tau_{\text{high-speed rail}}$. These results indicate that the decrease of average vehicle occupancy as well as the increment of the high-speed railway pricing will be good candidates for the control of traffic flows from Wuhan to those destinations with high incidence rate risks. Remember that reducing n will also lead to a decrease of Q_{auto}, which is against the original aim of promoting the aggregated auto flows. This kind of contradictive effect of the adjustment of a certain parameter should be taken account of and considerate measures should be conducted. With the aim of increasing Q_{auto} and decreasing $Q_{\text{high incidence}}$ at the same time, we find a compromise in the conflict mentioned above, which raises the prices of high-speed railway tickets by a large amount while making a small improvement on the average vehicle occupancy or keeping the average vehicle occupancy without any change.

4. Conclusion

In this paper, a nested logit-based multimodal traffic flow distribution model and a solution algorithm are proposed. The model is designed taking account of experiences learned from historical data as well as making use of information collected from the real transportation system. The proposed model is verified by the application to a real-life problem of the demand distribution from Wuhan to other nationwide regions during the outbreak of COVID-19. The estimation results in the case show that the model proposed in this work delivers a desirable performance on demand distribution estimation. The results of the estimation of the number of incidence cases reveal that the spread of the epidemic is not linear with respect to the estimated traffic flow distribution results. And further analysis on this result inspires us that the

spread of the crisis is not purely dependent on the transportation situation, but also affected on the one hand by the control methods conducted by the public power and on the other hand by the frequency of local economic activities as well as the occurrence number of crowd-collected activities. The analysis of the role of lockdown on different travel modes reflects that lockdown on the high-speed railway has the most prominent impact on the traffic flow increment of other travel modes, and a lockdown on a certain travel mode causes different extent of aggregated mode flow increment of other travel modes. It is important to measure the magnitude of correlation between lockdown on a certain travel mode and the traffic flow increase of other travel modes. The public transport mode which has a high correlation with the lockdown policy needs intensified management to prevent virus from spreading through transportation. Furthermore, sensitivity analysis is implemented in this study, and based on the results of which, we work out a compromise solution for stimulating the traffic flow of automobile and reducing the demands of target regions with high incidence rates at the same time.

Data Availability

The data used to support this study are available at Tencent social network's 2017 Spring Festival geographic positioning data platform, Baidu Migration Big Data Platform (https://qianxi.baidu.com/2020/), and https://news.qq.com/zt2020/page/feiyan.htm#/global.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] R. M. Anderson, R. M. May, and B. Anderson, "The mathematics of infection. (Book reviews: infectious diseases of humans. dynamics and control)," *Science*, vol. 254, no. 2, pp. 591-592, 1991.

[2] N. Bailey, "The mathematical theory of infectious diseases and its applications," *Immunology*, vol. 34, no. 5, pp. 955-956, 1978.

[3] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Review*, vol. 42, no. 4, pp. 599–653, 2000.

[4] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

[5] C. Li, W. Hu, and T. Huang, "Stability and bifurcation analysis of a modified epidemic model for computer viruses," *Mathematical Problems in Engineering*, vol. 2014, Article ID 784684, 14 pages, 2014.

[6] C. Fraser, C. A. Donnelly, S. Cauchemez et al., "Pandemic potential of a strain of influenza A (H1N1): early findings," *Science*, vol. 324, no. 5934, pp. 1557–1561, 2009.

[7] M. E. J. Newman, S. Forrest, and J. Balthrop, "Email networks and the spread of computer viruses," *Physical Review E*, vol. 66, no. 3, Article ID 035101, 2002.

[8] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics in finite size scale-free networks," *Physical Review E*, vol. 65, no. 3, Article ID 035108, 2002.

[9] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: an eigenvalue viewpoint," in *Proceedings of the 22nd International Symposium on Reliable Distributed Systems (SRDS'03)*, pp. 25–34, IEEE, Florence, Italy, October 2003.

[10] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, "Epidemic thresholds in real networks," *ACM Transactions on Information and System Security*, vol. 10, no. 4, pp. 1311–1326, 2008.

[11] I. Otero-Muras, G. Szederkényi, K. M. Hangos, and A. A. Alonso, "Dynamic analysis and control of biochemical reaction networks," *Mathematics and Computers in Simulation*, vol. 79, no. 4, pp. 999–1009, 2008.

[12] G. Pang and L. Chen, "Dynamic analysis of a pest-epidemic model with impulsive control," *Mathematics and Computers in Simulation*, vol. 79, no. 1, pp. 72–84, 2008.

[13] C. Jin and X.-Y. Wang, "Analysis and control stratagems of flash disk virus dynamic propagation model," *Security and Communication Networks*, vol. 5, no. 2, pp. 226–235, 2012.

[14] X. Y. Huang, C. Jin, and S. L. Jin, "Dynamic propagation model of blue-tooth virus on smart phones," *Advanced Materials Research*, vol. 121-122, pp. 620–626, 2010.

[15] Y. H. Li, J. X. Pan, and Z. Jin, "Dynamic modeling and analysis of the email virus propagation," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 472072, 22 pages, 2012.

[16] M. Jackson and B. M. Chen-Charpentier, "A model of biological control of plant virus propagation with delays," *Journal of Computational and Applied Mathematics*, vol. 330, pp. 855–865, 2017.

[17] F. Jia and G. Lv, "Dynamic analysis of a stochastic rumor propagation model," *Physica A: Statistical Mechanics and Its Applications*, vol. 490, pp. 613–623, 2018.

[18] J. Zhang, J. Jiang, Y. Chen, K. Yang, H. L. You, and Y. W. Chen, "Epidemic spreading characteristics and immunity measures based on complex network with contact strength and community structure," *Mathematical Problems in Engineering*, vol. 2015, Article ID 316092, 10 pages, 2015.

[19] S. Kitthamkesorn and A. Chen, "A path-size weibit stochastic user equilibrium model," *Transportation Research Part B: Methodological*, vol. 57, no. 11, pp. 378–397, 2013.

[20] C. Yang and A. Chen, "Sensitivity analysis of the combined travel demand model with applications," *European Journal of Operational Research*, vol. 198, no. 3, pp. 909–921, 2009.

[21] C. Yang, A. Chen, X. Xu, and S. C. Wong, "Sensitivity-based uncertainty analysis of a combined travel demand model," *Transportation Research Part B: Methodological*, vol. 57, pp. 225–244, 2013.

WILEY | Hindawi

Research Article

A Fissile Ripple Spreading Algorithm to Solve Time-Dependent Vehicle Routing Problem via Coevolutionary Path Optimization

Wen Xu [ID] **and JiaJun Li**

School of Management, Northwestern Polytechnical University, Xi'an 710072, China

Correspondence should be addressed to Wen Xu; 18078115229@126.com

The time-dependent vehicle routing problems have lately received great attention for logistics companies due to their crucial roles in reducing the time and economic costs, as well as fuel consumption and carbon emissions. However, the dynamic routing environment and traffic congestions have made it challenging to make the actual travelling trajectory optimal during the delivery process. To overcome this challenge, this study proposed an unconventional path optimization approach, fissile ripple spreading algorithm (FRSA), which is based on the advanced structure of coevolutionary path optimization (CEPO). The objective of the proposed model is to minimize the travelling time and path length of the vehicle, which are the popular indicators in path optimization. Some significant factors usually ignored in other research are considered in this study, such as congestion evolution, routing environment dynamics, signal control, and the complicated correlation between delivery sequence and the shortest path. The effectiveness of the proposed approach was demonstrated well in two sets of simulated experiments. The results prove that the proposed FRSA can scientifically find out the optimal delivery trajectory in a single run via global research, effectively avoid traffic congestion, and decrease the total delivery costs. This finding paves a new way to explore a promising methodology for addressing the delivery sequence and the shortest path problems at the same time. This study can provide theoretical support for the practical application in logistics delivery.

1. Introduction

The vehicle routing problem (VRP) and its extended variants have recently received great attention in many fields, which are required to improve delivery efficiency and reduce the delivery costs of logistics companies [1–6]. In its basic form, some essential elements are often used to construct the foundation of the VRP resolution, for instance, several depots and a set of customers [7]. Each customer is located in a different geographic position and demands a specific amount of goods, which are carried to the designative destinations by a fleet of vehicles [8]. The vehicles will depart from the depots and finally return to the same depots, if the given delivery assignments are completed. The typical constraints of VRP are to pose a strict requirement on the vehicle travelling trajectory. That is, all matched sites can be visited only once along with the planned vehicle route. The

common optimization objective of VRP is to seek the optimal delivery sequence and the shortest path performed by the specific vehicle. It requires that all customers can receive the associated goods, aiming at cutting down the related travelling time and path length of the overall routes.

Due to the boosting number of vehicles in urban networks, regular traffic congestion has been a common phenomenon that imposes some negative impacts on social development and daily life [9]. Especially for the logistics industry, several inevitable challenges are brought to impede the further research of the VRP and its multiple variants. How to effectively solve VRP plays a considerable role in minimizing delivery costs and ecological environment protection [10]. For example, the daily traffic congestions may cause huge uncertainties for vehicle route planning, when the vehicles intend to travel across a planned path with shorter length but will waste a long time waiting at the

blocked areas [9]. The fuel consumption highly depends on the length and travelling time of the route. Hence, traffic congestion tends to result in high fuel consumption and air pollution with CO_2 emissions because the engine must keep running during waiting periods. According to reliable data reported by the International Energy Agency, the transportation sector has contributed to the second-largest CO_2 emissions in 2015 [11]. In recent decades, how to reduce carbon emissions and fuel consumption has become a significant issue that should be paid more attention to VRP [12–16]. Under the pressure of additional costs and environmental protection, the current concerns of logistics companies are to exploit an appropriate approach for addressing the VRP [17–19].

Most of the research studies investigated the VRP proposed some ideal hypotheses. For example, the vehicle travel speed between two nodes is assumed to be fixed, which means that the travelling time is simply regarded as a constant value. Under the background of the time-dependent vehicle routing problem (TDVRP), it is impractical to be supposed in the real world because the vehicle speeds may vary with the changes of traffic condition. In other words, the travel speed between two nodes is time-dependent and usually relies on the starting time. When the vehicle reaches a node, the related link and node that will be visited soon may transform into blocked from the accessible state, influenced by sudden traffic congestion. As a result, an extra detour strategy or waiting behaviour both make the travelling time very long, even though the planned path possesses the shortest length. The longer path between two sites may not experience traffic congestion during the same period. A long path is not consistent with large travelling time [20]. Therefore, the vehicle route should take into account the changes in traffic conditions. Malandraki was the first to investigate TDVRP via exploring the impacts of traffic congestion on vehicle speeds [21]. Eglese et al. designed a timetable of the shortest paths based on a practical network, to present the expected route distance and travel time between two nodes [22]. Donati et al. developed an advanced multi-ant-colony system to minimize the total travel time, during a fleet of vehicles travelling at time-dependent speeds with a fixed capacity [23]. Kok et al. employed a modified Dijkstra algorithm and dynamic programming heuristic to implement an appropriate strategy to avoid traffic congestion in TDVRP [24]. Huang et al. developed the TDVRP with path flexibility (TDVRPPF) under the stochastic traffic conditions and gave the path decisions that relied on the departure time [25]. Kuo adopted simulated annealing to minimize the fuel consumption used in the route for TDVRP [26]. Ma et al. proposed a combined order selection to address the TDVRP with time windows for perishable product delivery [27]. Taş et al. studied the TDVRP considering soft time window and stochastic travel time, and applied tabu search and adaptive large neighbourhood search to prove the effectiveness of the proposed model [28]. Andres Figliozzi proposed a promising algorithm that can well solve TDVRP with a soft or hard time window without any alteration [29].

To address TDVRP in the real world, this study aims to minimize the travelling time and path length of the route by solving the optimal delivery sequence and the best path at the same time. A novel fissile ripple spreading algorithm (FRSA) was proposed to find the optimal solution by avoiding daily traffic congestion in advance, or waiting to pass through a rapidly shrinking congestion. The coevolutionary path optimization (CEPO) technology was introduced to apply in the proposed method, according to the significant characteristic of traffic operational rules and delivery demand. The effectiveness of the proposed approach was illustrated in two sets of simulation experiments. The results proved that the proposed FRSA can effectively reduce the travelling time and path length of the vehicle route, which can decline fuel consumption and avoid regular traffic congestion.

The rest of this study is organized as follows: Section 2 explains the problem description and the details of CEPO technology introduced in this work. Section 3 presents the basic idea of the proposed FRSA and shows how to integrate CEPO into it. Section 4 shows the experimental results by comparing the performance of FRSA with another compound method. Section 5 discusses some possible reasons for the outstanding results derived from FRSA and indicates its potential application. Section 6 outlines the conclusions of this work.

2. Coevolutionary Path Optimization

2.1. Problem Description. As a critical solution to reduce travel time and fuel consumption, TDVRP has been studied in extensive fields, which provides a great potential alternative to boost delivery efficiency and decrease economic cost. However, most of the existing studies ignored the practical network topology and separate the essential relationship between the shortest path of two sites and delivery sequence. Generally speaking, the delivery sequence decision is greatly affected by the weight values between two sites, which are obtained from the shortest path optimization. On the other hand, if a vehicle set out from a site at different periods, it tends to encounter different traffic conditions, and the shortest paths are likely to be changed. One can see that the path optimization results are also influenced by the delivery sequence. Furthermore, few studies focus on how to design a general algorithm to consider traffic congestion. Traffic congestion is the most common phenomenon in urban roads, which tends to cause a lot of time delay for vehicles as well as fuel consumption, along with some potential damage for emergency rescue. Therefore, although some encouraging progress has been made, developing a better intelligent algorithm with excellent optimization performance and good applicability is less successful and still requires further research. To overcome these research gaps, this paper firstly proposed a universal approach called FRSA to address TDVRP, considering dynamic traffic congestion and the coupling relationship between the shortest path and delivery sequence.

Thanks to the foundation offered by the previous research, the proposed TDVRP in this study aims to make the

actual delivery trajectory optimal, and it can be separated into two important issues. The one is to propose an appropriate path optimization method to find out the shortest path between two sites. The other is to exploit a feasible algorithm to determine the best delivery sequence with small travelling time and short path length on account of the path optimization method. In other words, the goal is to minimize the total travel time and path length by optimizing the vehicle delivery order and the delivery paths between any two sites.

According to the formal description of TDVRP in previous studies, several crucial rules should be established. A series of predetermined sites must be reached by a specific vehicle that is located in a single depot. The locations and the demands of each site are known, while the location and capacity of the related depot are given as well. Due to the complicated effects of traffic flow and road geometry, vehicle speeds may vary over time in a road section [30–33]. Additionally, when the vehicle travels on different roads or different driving directions, it tends to encounter various speeds and traffic conditions. To clarify the application scope of this study, two assumptions must be made. (1) The vehicle is allowed to wait at signalized intersections or congested links, and the delay caused by this inevitable procedure will be contained in total travel time. (2) The vehicle can get access to future traffic data from the traffic information prediction platform or big data prediction algorithm. The future traffic data include congestion scope, the average delay of signalized intersections, and average traffic speeds of all roads. (3) Any change in the routing environment will eventually be reflected by the certain variation of traffic speeds in related roads.

2.2. Basic Idea of CEPO.

In this paper, an advanced framework of CEPO is introduced to improve the path optimization algorithm, which concentrates on how to find out the best path between different sites. As a newly emerging approach, CEPO is firstly proposed by Hu et al. who have certified its outstanding performance and feasibility in discrete urban networks [34]. Moreover, another advanced variant timing coevolutionary path optimization (TCEPO) was developed to prove the effectiveness of CEPO by several comparative simulation experiments [35]. Due to its coevolutionary optimization mechanism, it can achieve a global optimal result as long as the given prediction data are precise enough. The prediction data in this study are obtained from those intelligent prediction algorithms, which are used to predict the short-term traffic flow or traffic speed based on historical traffic data. The prediction algorithms tend to adopt machine learning theory to predict traffic dynamics and usually achieve more than 90% of mean prediction accuracy [36–38]. Some of them can achieve higher prediction accuracy, even more than 95% [38]. These encouraging results can provide the predictive traffic data, which are assumption basis for the proposed method. To explain how it can implement a superior optimization process, the basic idea of CEPO is illustrated in Figure 1.
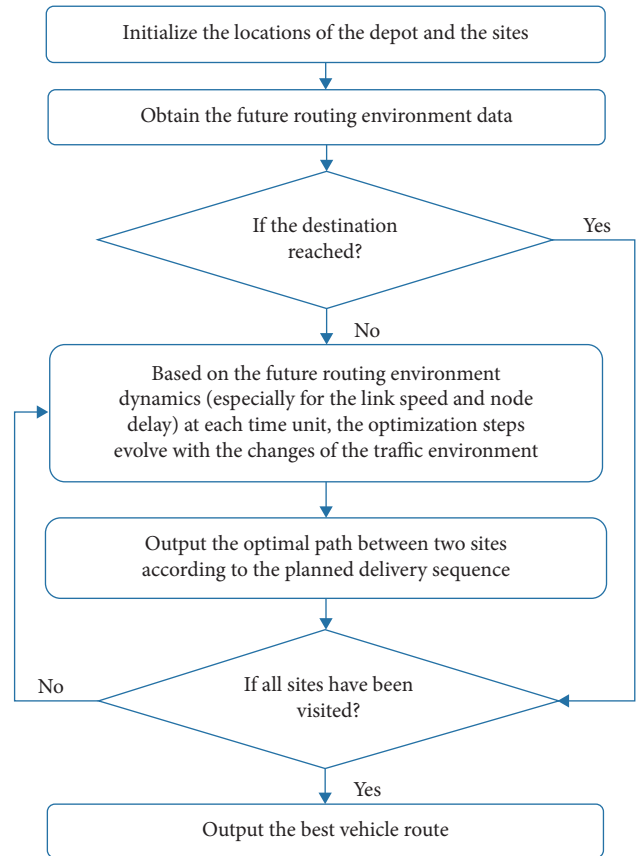
Figure 1: The framework of the optimization process for the CEPO technology.

The definition of CEPO is, under the predicted traffic environment dynamics, each time-unit-oriented optimization step coevolves with the traffic speed variation of associated links once the traffic environment changes in routing networks, to make actual travelling trajectory optimal in a single run. All optimization steps will be carried out before the vehicle sets out from the depot. Unlike other mainstream research studies based on online reoptimization (OLRO), CEPO has more advantages in utilizing future traffic information and optimization accuracy. This is because the future routing environment dynamics are considered into the optimization procedure, which brings an extra superiority to know the evolutionary trend of routing environment in advance. As a result, it can help a vehicle to avoid wrong detour decisions or escape from congestion areas in advance. No matter which kind of roles it plays, CEPO poses a positive effect on the best route decisions that can help vehicle travel across the temporary congestion areas as fast as possible.

2.3. Mathematic Expression of CEPO.

It is assumed that the predictable routing environment dynamics and vehicle delivery sequence are both given, CEPO aims to address the minimization problem in terms of time cost. Travelling time is usually a major concern in path optimization, particularly in the dynamic urban routing environment [39, 40]. Moreover,

as explained in Section 2.3, CEPO technology is a time-unit-oriented optimization methodology. Thus, it is an effective way to use travelling time as the optimization objective of CEPO. The objective function can be established as follows:

$$\min F_C = \sum_{k=1}^{n+1} f_k(P_k, L(P_k)), \tag{1}$$

which is subject to the following equations:

$$\begin{aligned} f_k(P_k, P_k(i+1)) &= D_t(P_k(i)) \\ &+ C_t(P_k(i), P_k(i+1)), \quad i = 1, \ldots, L(P_k) - 1, \end{aligned} \tag{2}$$

$$f_k(P_k, 1) = 0, \tag{3}$$

where F_c is the function calculating the time cost of all optimization paths. n denotes the number of delivery sites. P_k denotes the kth optimization path between two sites in the presupposed delivery sequence. $L(P_k)$ denotes the number of all nodes in path P_k. $f_k(P_k, L(P_k))$ denotes the time to travel through all nodes in path P_k. $f_k(P_k, P_k(i+1))$ denotes the time to pass through the node $P_k(i+1)$ when the vehicle just reaches the node $P_k(i)$. $D_t(P_k(i))$ denotes the time delay to cross over the node $P_k(i)$ in path P_k at time t. $C_t(P_k(i), P_k(i+1))$ denotes the travelling time associated with the link between nodes $P_k(i)$ and $P_k(i+1)$ at time t.

In this section, the objective function in (1) is defined as minimizing the total travelling time of the complete delivery loop. Equations (2) and (3) clearly show the calculation derivation of the travelling time, consisting of the time costs that pass through the nodes and links, respectively. It is believed that the time cost of a subpath in P_k tends to be dependent on when the subpath is to be travelled through, which is often overlooked in other studies. Thus, if the routing environment dynamics are predictable, one-off optimization will be prospectively conducted to seek the best path P^* in the time-varying traffic environment before the vehicle starts to perform the delivery assignments. This means that the reoptimization of the path P^* required in OLRO-based methods is unnecessary to use at each time unit.

3. Fissile Ripple Spreading Algorithm Based on Global Search

3.1. Algorithm Origin and Development. To realize the CEPO method as mentioned in Section 2, it is essential to exploit an appropriate algorithm to adapt to the synchronous state changes of nodes and links. It has achieved a great success for deterministic and heuristic methods in solving static path optimization and time-dependent path optimization. For example, the Dijkstra algorithm, A* algorithm, genetic algorithm, particle swarm optimization algorithm, simulated annealing algorithm, and ant colony algorithm. It is unlikely to apply them to an effective realization of CEPO. One important reason is that their basic computational step is the link-oriented analysis that associates with a long period, which may cause serious scalability problems [30]. What makes things worse is that

it is impossible to pre-know which link will be accessed or when the given links will be accessed. Additionally, it must be considered for the improvement of delivery strategy and the best-first search of the shortest paths between sites. According to the latest knowledge, it seems to be quite hard to synchronously perform this procedure for most methods. This is why in a time-dependent routing environment, an advanced optimization algorithm needs to be designed to achieve the CEPO, to cover all necessary changes in the route network.

FRSA is an evolutionary branch of the ripple spreading algorithm (RSA), which is first reported by Hu and Liao in 2016 [41]. Like other evolutionary computation algorithms, the basic idea of FRSA takes inspiration from nature and intends to imitate the ripple spreading phenomenon. The core technology of FRSA presents excellent optimality and surprisingly robustness in resolving complex problems, which is conducted by a multilevel and fissionable ripple relay race. FRSA is an agent-based, deterministic, and discrete simulation model, which provides a rare opportunity to evolve with the changes in the predicted environment at each time unit. During a single run of FRSA, it delivers a great number of ripples and the spreading speeds of ripples reflected in the changes of links to program optimization steps. Furthermore, when a ripple activates a site from the last site, the fissile mechanism is permitted to carry out more independent ripple relay races that explore optimal resolution in global research. Thus, once a single run of FRSA has been done, optimal delivery sequence and the shortest paths of given OD pairs (site to site or depot to site) can be obtained.

3.2. Fissile Ripple Relay Race Based on Global Research. To realize the CEPO methodology, the optimization principle of the FRSA is extended to be divided into two parts, that is, independent ripple relay race and fissile program. An original ripple relay race is applied in a static route network [42]. But in the actual application, especially for the urban network, traffic environment dynamics should be integrated into an independent ripple relay. The illustration of an independent ripple relay race is given in Figure 2.

In the ripple relay race, the initial ripple is activated to split into several ripples from the origin node, and new ripples travel toward the adjacent nodes along with the connected links. When a ripple reaches an unvisited node, it will trigger new ripples at that node, and the unvisited state will be transformed into an active state. Those new ripples will continue to spread out until arriving in the adjacent nodes, while their spreading speeds are consistent with the traffic speeds of the link. As the ripple relay race goes on, numerous ripples will be generated to compete with each other until a ripple reaches the given destination node. Eventually, the shortest path from the origin to the destination will be determined by tracking back the visiting node order, when the ripple reaches the destination firstly. Therefore, the travelling time of the shortest path between sites can be calculated as the end of an independent ripple relay race.
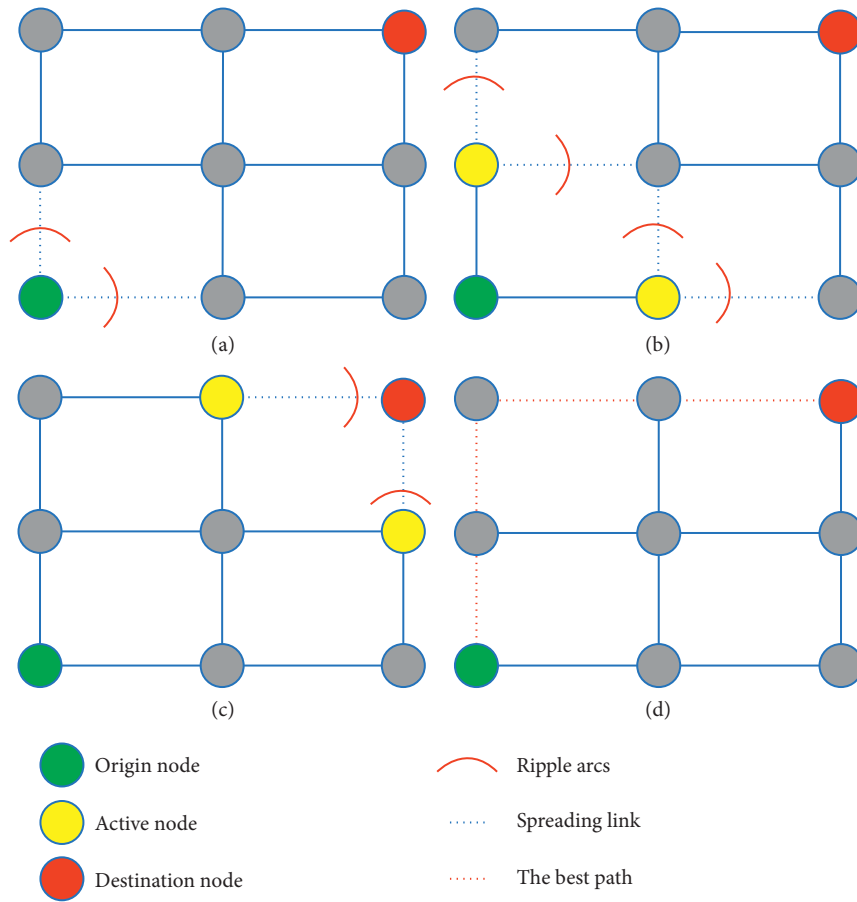
Figure 2: The illustration of an independent ripple relay race.

Although the independent relay race just offers a supportive foundation for addressing TDVRP in a dynamic routing environment, it is still less successful. It seems to be an ongoing challenge that how to decide an optimal delivery sequence with the smallest travel time cost. Hence, how to apply ripple relay race to resolve the best visiting order for all preset sites may be a question, which is worth thoroughly pondering. However, thanks to the development of computer science technologies, stronger computational ability can be used to execute big data processing. It opens a door to the fissile program that can evolve more independent relay races, to search more loop paths associated with different visiting orders.

To explore the optimal delivery sequence, it is a critical way to search for all delivery schemes in parallel by performing extensive ripple relay races, which are independent of each other. The diagram of the fissile program in FRSA is given in Figure 3. When the nearest site will be triggered by a ripple from the depot, it will generate a unique list recording activation order of the site. At the same time, a new ripple relay race starts to run at the site. All ripples from the new ripple relay race will not be affected by any ripple from the previous one. In other words, the ripples of different relay races are independent of each other, although they may arrive in the same nodes or links simultaneously. All ripple relay races run in the original size network. What is more
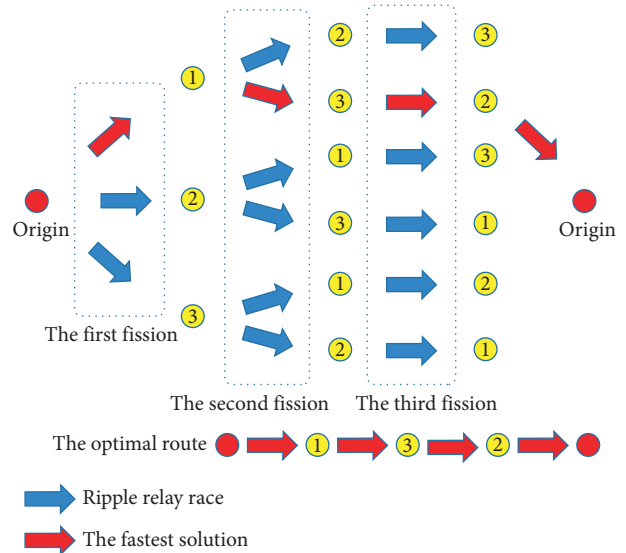
Figure 3: The schematic diagram of the fissile program in FRSA.

important, the previous ripple relay race will continue to program, even if it has given rise to a new ripple relay race at the unvisited site. The other ripples from the previous relay race still own similar ability to trigger other sites, as long as these sites are not contained in the list of activity orders

recorded by the associated ripples. This strict rule of FRSA indicates that a site will not be triggered twice in the same ripple relay race but remains accessible to others. As time goes on, more and more independent ripple relay races are conducted to evolve with the routing environment dynamics. The optimization of FRSA will be terminated, when all sites are recorded in the list of activation order and the related ripple finally reaches the depot in the first place. The optimal loop path can be gained by a single run of FRSA after a large number of computations for time-unit-oriented analysis.

3.3. Realization of FRSA considering the CEPO Technology.
To make the actual loop path optimal, it is crucial to permit ripples to wait at a node that represents a signalized interaction in practical application. The waiting behaviour of ripples reflects that the vehicles have to obey legal traffic rules, such as the restriction of red light and forward vehicles. Another condition is that the ripples can wait at a link or a node when they are temporarily blocked due to the evolution of the congestion area. However, this waiting behaviour provides a potential likelihood of waiting at a link or node until they become accessible again, which may take less time than travelling around the block area. It can impose a positive impact on avoiding the wrong detour, which is helpful to find out the shortest path with the smallest travelling time. Therefore, to extend the FRSA to realize the CEPO technology, it needs to introduce a waiting behaviour of ripple into the relay race, to imitate the real performing actions of vehicles in the urban roads.

To understand the outline of FRSA, some variables need to be calibrated to explain how it works. Let N denote the total number of nodes in the network; n denotes the number of sites. For the sake of simplicity but without losing generality, this paper supposes that there are one hundred nodes, three sites, and one depot in the network. Let node 1 represent depot. m is the number of adjacent nodes for any node and sets $m = 4$, which means each node is connected to four nodes. In a ripple relay race, node i can be triggered with unlimited times and its ripple is called ripple i, which can split into several arcs spreading to adjacent nodes. $S_r(i)$ is the state of the ripple i, $S_r(i) = 1, 2, 3$ means ripple i remains inactive, active, and waiting, respectively. $r(i, j)$ is the radius of ripple arcs from node i to node j ($i \neq j$). $V_t(i, j)$ denotes the spreading speed of ripple arc from node i to node j at time t, which is related to expand the associated arcs radius $r(i, j)$. $F_r(i)$ denotes a crucial variable that records which ripple triggered the node i. For instance, $F_r(i) = j$ means node i is triggered by the ripple from node j, which offers detailed information to track back the travelling trajectory. When the ripple relay races are terminated, the optimal loop path P^* will be determined by iterative deduction of $F_r(i)$, and the expression is illustrated in the following equation:

$$P^*(i) = \begin{cases} 1, & \text{if } i = L(P^*), \\ F_r(P^*(i+1)), & \text{if } i < L(P^*). \end{cases} \quad (4)$$

To avoid the repeated activation of the sites in the same ripple relay race, $M_r(i)$ is used to record the existing travelling trajectory of the ripple i via tracking back $F_r(i)$. One can see that any site will not be repeated on each $M_r(i)$ except for the depot, which is the origin as well as the destination. In the real world, the spreading speed of ripple i is subject to the speed limit of the urban roads. Thus, $V_t(i, j)$ must satisfy the following constraints, which are given in the following equations:

$$0 < V_t(i, j) \leq V_{\lim}(i, j), \quad (5)$$

$$V_t(i, j) \times t_{\text{unit}} \leq L(i, j), \quad (6)$$

where $V_{\lim}(i, j)$ is the maximal speed limit of the link (i, j). t_{unit} is the time unit that is small enough to capture the major dynamics of the routing environment. $L(i, j)$ is the length of the link (i, j). As time goes on, the arc radius of the active ripple will be updated according to the associated spreading. The calculation formula of the arc radius is shown in the following equation:

$$r(i, j) = r(i, j) + V_t(i, j) \times t_{\text{unit}}. \quad (7)$$

If the node i is reached by the ripple j before the ending of a time unit, the new ripple i would continue to spread out in the remaining time t_{re}. The spreading radius of arcs from the ripple i to its adjacent node k in the remaining time t_{re} can be obtained according to the following equations:

$$t_{\text{re}} = \frac{r(j, i) - L(j, i)}{V_t(j, i)}, \quad (8)$$

$$r(i, k) = r(i, k) + V_t(i, k) \times t_{\text{re}}, \quad (9)$$

where t_{re} is the remaining time in a time unit when a ripple arrives at a node. $r(i, k)$ denotes the spreading radius of arcs from the ripple i to its adjacent node k in the remaining time t_{re}.

The flowchart of the FRSA based on the CEPO methodology is demonstrated in Figure 4. One can see that Step 1 to Step 3 well describe how to construct a simulation network and initialize some parameters as the prior preparation. Step 4 defines the termination condition of the ripple relay race; that is, all sites are visited by a specific ripple that finally returns to the depot. Step 5 indicates the function that converts the state of the node from waiting into active if the node is accessible. Step 6 gives the derivation formula of the arc radius for an active ripple. Step 7 means that the remaining time will be used to develop the arc radius of the new active ripple, which is triggered by other ripples before the ending of a time unit. Step 8 demonstrates the requirement of a waiting behaviour at a triggered node, while it is blocked. Step 9 updates the state of the nodes to inactive again, when their all ripples have reached adjacent nodes. Step 10 judges whether FRSA needs to start a new ripple relay race from a node. Step 11 explains the mandatory operation rules for a new ripple relay race. Step 12 outputs the optimal travelling trajectory based on complete $M_r(1)$.

Step 1: create N nodes network; give the location and number of depots and all sites. Set node 1 as depot.

Step 2: initialize parameters. For all nodes, $i = 1: N$; let $F_r(i) = 0$, $S_r(i) = 1$, $r(i, j)$.

Step 3: start the ripple relay race. Set $t = 0$, $S_r(1) = 2$, $M_r(1) = \{1\}$.

Step 4: if all sites are recorded in $M_r(1)$?

Yes

No

Step 5: convert the state of the nodes. For $i = 1: N$, whether the node i is accessible, $S_r(i) = 3$ and $F_r(i) > 0$. If so, then change the state of node i from waiting to active. Update $S_r(i) = 2$.

Step 6: calculate the arc radii of the ripples. For $i = 1: N$, whether $S_r(i) = 2$. If so, then update all arc radii of the ripple i over time according to formula (7). Let $t = t + 1$.

Step 7: manage the remaining time. For $i = 1: N$, whether $S_r(i) = 1$ and node i is triggered by adjacent node j. If so, then update all arc radii of the ripple i in the remaining time according to formulas (8) and (9). Let $F_r(i) = j$ and $S_r(i) = 2$.

Step 8: realize the waiting behavior. For $i = 1: N$, whether the node i is triggered by its adjacent ripple j but blocked, and $S_r(i) = 0$. If so, then let $F_r(i) = j$, $r(i) = 0$, and $S_r(i) = 3$, where $r(i)$ denotes all arc radii of ripple i.

Step 9: update the state of the ripple. For $i = 1: N$, whether the ripple i has triggered each adjacent node and $S_r(i) = 2$. If so, then let $S_r(i) = 1$.

Step 10: if any site is reached by the ripple?

No

Yes

Step 11: generate new independent ripple relay race. For any ripple, i will not trigger the sites that are contained in $M_r(i)$ and the depot except for all sites contained in $M_r(i)$ but still can trigger other normal nodes. Assuming that rhe site k is reached by the ripple j, then let $F_r(k) = j$, $S_r(k) = 2$, $M_r(k) = \{l, \dots, j, k\}$.

Step 12: output the optimal loop path P^*, which is recorded in $M_r(1)$.

FIGURE 4: The flowchart of the FRSA based on the CEPO methodology.

4. Experimental Results and Analysis

4.1. Static Delivery Sequence Determination Tests Compared with GA. In this experiment, to achieve a fair understanding of the performance of FRSA, four tests with different site numbers are conducted to resolve the traditional VRP. A widely used GA is selected to output the results for the comparative analysis. A series of sites (red points) and depot (green point) are generated as follows: the location of the depot is set as the left side as an example and the locations of the sites are randomly distributed in the right district. All nodes have the associated links and connected to their

adjacent nodes, and the weights of the links depend on their length. To show the generality of the proposed FRSA in the delivery application scenario, the depot is given only one, while there are four sites scales with $n = 10$, 12, 14, and 16, respectively. The diagrams of the node connection for all tests are shown in Figure 5.

Since the ripple relay race of FRSA is the time-unit-oriented process, it will be terminated once a ripple firstly finds the shortest path in a certain time unit. After that, the FRSA will arrange all the complete routes in that time unit, according to the path length from small to large to select the optimal path. For FRSA, an appropriate spreading speed can
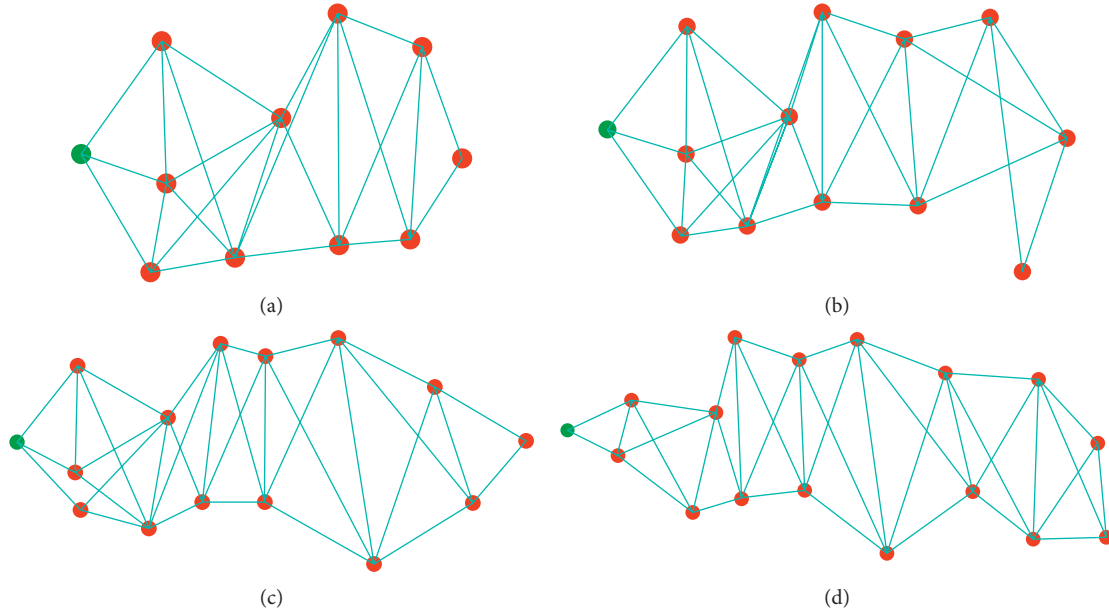
FIGURE 5: The diagrams of the node connection in different network cases. (a) $n = 10$, (b) $n = 12$, (c) $n = 14$, and (d) $n = 16$.

improve the computational efficiency and reduce computation or storage pressures. To ensure the computation efficiency of FRSA optimal, the spreading speed of ripples cannot be too large or too small in the network. If the spreading speed is relatively high, the path length travelled by ripples will be pretty long in a time unit. When the optimal route is solved in a time unit, a great number of travelling trajectories that are closed to the optimal path are likely to be found in the remaining time and still need to be arranged. The unexpected results pose extra work on the calculation module, which means that the computational time required for sorting path length of each complete route increases accordingly. If the spreading speed is relatively low, resulting in more time units required to extend the arc radius of the ripples. The workload of updating the ripple state significantly increases with the growth of time units used in a single run. Hence, the ripple spreading speed can neither be too large nor too small. After comparing multiple spreading speeds through simulated tests, this paper proves that selecting 60 km/h has better computing efficiency. Additionally, this speed also conforms to the speed limit management regulations of urban roads in China. As a result, this experiment sets ripple spreading speed $v(i, j) = 60$ km/h.

The related results of FRSA and GA are given in Table 1 and Figure 6. One can see that, in terms of PL, it is apparent to observe that FRSA still outperforms GA in all kinds of nodes sizes. The distance of the optimized paths given by FRSA is approximately 33.5% to 57.7% smaller than that of GA. Compared with GA, the reduction rate of FRSA reaches the minimal value at $n = 12$ and peaks at $n = 14$ in terms of PL. One possible reason is that the core technology of FRSA is to globally explore all paths, and eventually ended when the ripple has visited all sites without repetitive activation and returns to the origin. Compared with GA, it

TABLE 1: The optimization results of FRSA and GA in different nodes size scenarios.

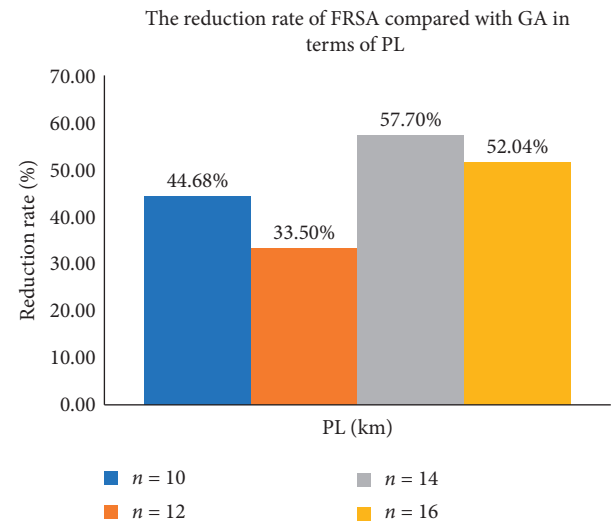|		$n = 10$	$n = 12$	$n = 14$	$n = 16$
PL (km)	FRSA	**33.17**	**51.9**	**45.96**	**53.97**
	GA	59.96	78.05	108.66	112.53
CT (s)	FRSA	**2.63**	**75.07**	**128.45**	**1062.82**
	GA	0.1	0.11	0.11	0.16

FIGURE 6: The reduction rate of the optimal paths offered by FRSA compared with GA in terms of PL.

possesses better search capacity and will not fall into a local optimum. Unfortunately, as a sacrifice of the huge advantage in solving the best path, the CT of FRSA remarkably goes up with the expansion of node sizes.

However, unlike other global searching algorithms, FRSA needs no requirements to traversal all results. This is because the optimization steps are divided into a lot of time units, and the ripple relay race will be terminated once the first group of paths is found at a certain time unit. It means that extensive data calculation used for searching other longer loop paths has been omitted. In this respect, FRSA has achieved great progress in terms of CT without losing global searching capacity, although it is relatively large compared with those heuristic algorithms.

4.2. Delivery Trajectory Optimization Tests considering Dynamic Routing Environment. As mentioned in Section 2.1, this study is especially interested in addressing a significant issue via designing an efficient methodology based on CEPO technology, which aims to make the actual delivery trajectory optimal. What makes things difficult is that the simultaneous resolution of the shortest paths between sites and the optimal delivery sequence poses a great challenge on the current methods and tends to be ignored. Therefore, this study investigates the interaction relationship of them and tests the performance of FRSA based on CEPO, comparing to the mainstream heuristic algorithm based on OLRO.

Different from Section 4.1, traffic dynamics of networks are introduced to imitate the actual traffic environment. However, travelling time is significantly affected by traffic flow in urban roads. When the traffic flow of the specific road section is large, the associated vehicle density also is high, which means the space headway between vehicles will be small. However, the smaller space headway between vehicles, along with the more serious disturbance effect between vehicles. This is because the drivers are subject to the front vehicles or the lane-changing behaviour of the adjacent vehicles. The frequency and severity of traffic flow turbulence will be significantly enlarged as the growth of traffic flow, such as dangerous deceleration behaviour, risky lane changes, and sudden accidents. Thus, the negative effects of traffic flow turbulence lead to the reduction of average traffic speed, which poses a direct impact on the increase of travelling time. Conversely, if the traffic flow on the road section is small, the speed of the vehicle tends to be less affected by that, and the travelling time is also reduced accordingly. Thus, travelling time also needs to be an additional assessment indicator, which is used to analyze the temporal efficiency of the optimized routes. In this study, a set of experiments is conducted to verify the promising application of the CEPO-based FRSA in terms of travelling time (TT), path length (PL), and computational time (CT).

To obtain the shortest path between sites in GA, a popular Dijkstra algorithm is introduced to achieve the realization of the OLRO [43–46]. In OLRO, the basic application of the Dijkstra algorithm is to repeat calculating the current shortest path based on the dynamic routing network in each time unit until the vehicle reaches the destination. Based on the current shortest paths between sites, GA is used to find out the optimal delivery sequence via iterative

computation and finally presents the complete delivery path. In CEPO, the proposed FRSA will give the optimal delivery trajectory in a single run according to the future routing environment.

In this part, different sizes of urban networks are randomly generated as following: the locations of N nodes are randomly evenly distributed in a rectangular area. There are 4 network scales with $N = 64, 144, 256,$ and 400, respectively. All nodes are connected to their neighbouring nodes, while each node has 3 or 4 connections. The shortest link is less than 150 m. The depot (green point) is always set as the central location of the network, while 8 sites (red points) are distributed around the depot randomly. Furthermore, to achieve the best computational efficiency, 1 min is selected to represent a time unit. However, the most important thing is to construct the routing environment dynamics, which can be described as traffic speed dynamics:

$$V_{t+1}(i, j) = V_t(i, j) + \sigma_t(i, j), \qquad (10)$$

where $\sigma_t(i, j)$ denotes the traffic speed variation of the link (i, j) at time t.

There are three hypotheses for processing nodes and links, as follows:

(i) To restore the actual traffic operation in the urban network, it is needed to determine the maximum speed and speed variation of the links at each time unit. Suppose $0 < V_t(i, j) \leq 80$ km/h and the absolute value of speed variation is less than 4 km/h for all links ($|\sigma_t(i, j)| \leq 4$) at each time unit.

(ii) Due to the compulsive traffic laws in the signalized intersections, the vehicles must wait at the nodes until they get access to the right to pass through. Thus, the stopping delay caused by waiting at nodes should be contained in travelling time during path optimization. Let $d(i)$ denote the average delay of node i, and its value randomly varies from 20 to 60 seconds.

(iii) As a common phenomenon in the urban network, the evolution of the congestion area also should be considered in TDVRP. As shown in Figure 7, the central nodes of congestion areas are evenly distributed in the network and have three kinds of states: spreading, static, and shrinking. It is supposed that the spreading or shrinking congestion areas will expand or narrow in every two minutes, and this action can evolve three times at most. Once the congestion areas spread to the maximum scope, it will keep unchanged in two minutes and then begin to shrink in a regular period.

Some experimental results are clearly given in Table 2 and Figure 8. Table 2 shows the results of the actual delivery trajectories given by FRSA and GA, in terms of TT, PL, and CT. One can see that no matter which scale of the route networks, the proposed FRSA (CEPO-based) outperforms GA (OLRO-based) in terms of TT and PL. From Figure 8, the TT of FRSA is approximately 11.86% to 20.25% smaller than that of GA. Moreover, the PL of FRSA is approximately
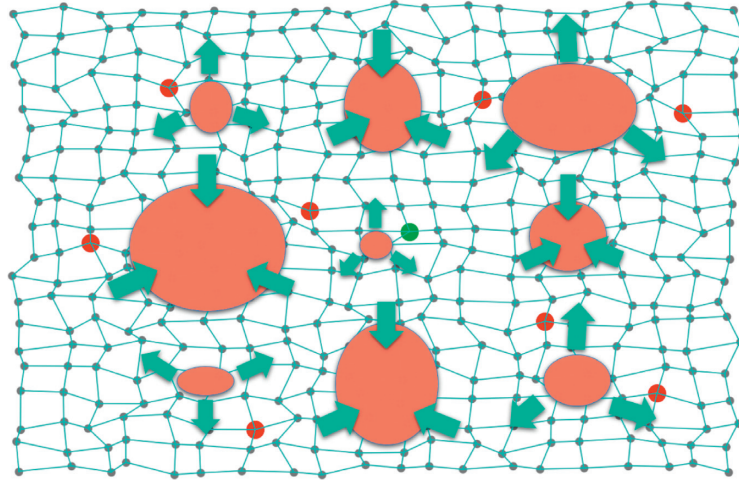
FIGURE 7: The schematic diagram of evolutionary dynamics for congestion areas ($N = 400$, $n = 8$).

TABLE 2: The optimized results of FRSA and GA in different dynamic networks.

		$N = 64$	$N = 144$	$N = 256$	$N = 400$
TT (min)	FRSA	**70.18**	**96.9**	**127.04**	**158.98**
	GA	83.85	121.51	146.97	180.37
PL (km)	FRSA	**39.82**	**46.83**	**73.42**	**92.11**
	GA	40.17	64.29	88.86	97.71
CT (s)	FRSA	**225.74**	**300.88**	**1463.46**	**6060.05**
	GA	441.45	1173.94	2269.65	6861.33

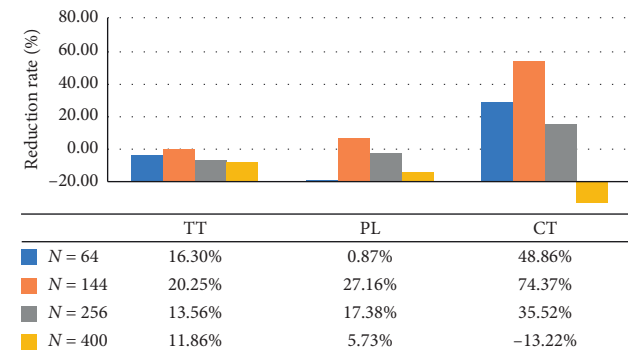|	TT	PL	CT
$N = 64$	16.30%	0.87%	48.86%
$N = 144$	20.25%	27.16%	74.37%
$N = 256$	13.56%	17.38%	35.52%
$N = 400$	11.86%	5.73%	−13.22%

FIGURE 8: The reduction rate of the optimal paths offered by FRSA compared with GA in terms of TT, PL, and CT.

0.86% to 27.16% shorter than that of GA. It is obvious to prove that the large PL is often inconsistent with the larger TT [34]. In small-size networks ($N = 64$, 144, and 256), the CT of FRSA is about 35.52% to 74.37% smaller than that of GA. Nevertheless, this trend tends to reverses when the network size becomes relatively large, and the CT of FRSA is 13.22% longer than that of GA when $N = 400$.

In terms of TT, PL, and CT, the reduction rate of FRSA gradually boosts from $n = 64$, then reaches the peak at $n = 144$, and finally declines smoothly over the boost of nodes. Compared with GA, surprisingly, the reduction percentage of FRSA is no-linear growth as the expansion of network size in terms of TT, PL, and CT.

5. Discussion

The main goal of this study was to validate the extensive applicability of FRSA based on the framework of CEPO. With this aim, this study performed traditional tests for FRSA and GA in the static networks, as presented in Section 4.1. For the sake of fairness, the reduction rate was used to measure the difference between FRSA and GA in terms of the TT, PL, and CT. The proposed FRSA exhibits superior searching ability and is far better than GA in terms of the cost of the related solutions. However, the encouraging findings also bring corresponding disadvantages. As regards the most serious defect, a massive amount of data is handled in FRSA, along with a sharp increase of CT as the number of sites goes on. An inspiring discovery thing is that, compared to other accurate approaches that achieved a similar effect, FRSA has cut down a tremendous number of computational works without seeking other possible solutions.

As discussed in the previous section, this study aims to propose a commonly used algorithm to address various popular branches of VRP, especially for TDVRP. Therefore, the time-varying traffic environment needs to be considered in testing FRSA as an inevitable development. As a crucial indicator, the TT of the solutions becomes a current focus. In a dynamic routing environment, FRSA is always superior to GA in terms of TT, PL, and CT, except for a special case ($N = 400$) that the CT of FRSA is larger than that of GA. There are three main reasons giving rise to the results that FRSA outperforms the GA in terms of TT and PL. One reason for this is that, in FRSA, the optimal delivery sequence and the shortest path between given OD pairs are solved effectively at the same time. GA just focuses on how to ensure the delivery sequence better, while Dijkstra only concentrates on how to find the best path between OD pairs. The second reason is that FRSA has more advantages to derive the optimal solution via global search, while many heuristic algorithms can only find an approximate solution through parallel computing. The last reason contributes to most to this result, the time-unit-oriented optimization steps of FRSA can coevolve with the changes of routing

environment and achieve the realization of CEPO. It owns unique features that avoid the coming congestion in advance, or allow waiting behaviour to prevent the wrong detour measure. Therefore, under the coupling effect of three critical factors, the actual travelling trajectory with a shorter TT can be obtained in the nature of things. Except for the big size network like $n = 400$, the CT of FRSA is smaller than that of GA in addressing TDVRP. As a result, a smaller CT is spent in FRSA to solve the path optimization in dynamic routing network. But, in $n = 400$ network, there are no more advantages for FRSA compared with GA. The main reason contributing to this is that each node provides an opportunity to trigger the ripples, and one node is usually triggered many times during optimization. More nodes also imply the boosting number of ripples will be generated, which asks for a larger CT to update the state transition and ripple spreading situation in a time unit.

However, FRSA possesses some distinct advantages in dealing with VRP, compared with the previous methods. The most outstanding one is that it shows a reliable guarantee of optimality in a time-varying routing environment. The other promising features, like a one-off run before a complicated trip begins and the positive avoidance of congestion areas, also proved that it is superior to the other. These results suggest that, in some urban networks (less than 400 nodes), it would be advisable to execute the FRSA to search the best solution as a practical implementation. As for large network scenarios, it needs to pay attention to whether there is a strict requirement for computational time. If not, FRSA will be a good choice to make the actual travelling trajectory better for addressing the existing VRP.

6. Conclusions

This work paves a new way to address various VRP variants, particularly for TDVRP whose main features have been studied for several decades. A mathematical model of FRSA is formulated based on the CEPO technology, which makes use of the future traffic environment to implement path optimization. Due to the effects of traffic flow, network structure, and driver characteristics, some critical factors like time-dependent traffic speeds, network topology, temporal congestion area evolution, different vehicle departure times, vehicle waiting behaviour, and network size are considered in the proposed model [47–49]. The objective of the proposed method is to minimize the sum of the delivery costs, such as travelling time and path length. Unlike the existing research studies, this study pays extra attention to explain the internal correlation between determining the delivery sequence and searching the shortest paths between sites, which tends to be ignored. However, FRSA owns several outstanding features, including a congestion avoidance approach, a single run optimization, waiting measure, parallel computing, coevolutionary planning, and global research. The effectiveness of the proposed approach was demonstrated in two sets of simulation experiments. The experimental results show the following: (1) FRSA presents an excellent improvement in reducing actual travelling time and path length, effectively avoiding the wrong detour and regular traffic congestions. It

can help to decrease the total distribution cost, fuel consumption, and carbon emissions, which can effectively protect the environment. (2) The optimal delivery sequence and the shortest paths can be derived simultaneously in the dynamic routing networks. This encouraging characteristic is significantly important and beneficial to make the actual travelling trajectory better, which is more likely to be adopted in the practical application compared with other methods. (3) The application of CEPO methodology considering the future routing environment dynamics will decline the travelling time. By contrast, other OLRO-based methods tend to take wrong coping strategies to avoid traffic congestions, which will increase the time and economic costs for logistics companies.

Despite some inspiring results that have been achieved, there still exist some limitations that need to be overcome. Further research may consider the extension of the proposed algorithmic framework to address other VRP variants in special scenarios, by taking into account time window, service time, vehicle capacity, site demand, and pickup. Additionally, the proposed FRSA can be further developed to resolve scalability problems and improve computational efficiency. As a new and effective deterministic algorithm, FRSA opens the door to the development of precise exploration.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] X. Jun-xiang and G. Jing-ni, "Research on logistics vehicle routing problem based on big data framework," *Journal of Transportation Systems Engineering and Information Technology*, vol. 18, no. 1, pp. 86–93, 2018.

[2] M. M. S. Abdulkader, Y. Gajpal, and T. Y. ElMekkawy, "Hybridized ant colony algorithm for the multi compartment vehicle routing problem," *Applied Soft Computing*, vol. 37, pp. 196–203, 2015.

[3] T. J. Ai and V. Kachitvichyanukul, "A particle swarm optimization for the vehicle routing problem with simultaneous pickup and delivery," *Computers & Operations Research*, vol. 36, no. 5, pp. 1693–1702, 2009.

[4] T. J. Ai and V. Kachitvichyanukul, "Particle swarm optimization and two solution representations for solving the capacitated vehicle routing problem," *Computers & Industrial Engineering*, vol. 56, no. 1, pp. 380–387, 2009.

[5] S. R. Ait Haddadene, N. Labadie, and C. Prodhon, "A GRASP × ILS for the vehicle routing problem with time

windows, synchronization and precedence constraints," *Expert Systems with Applications*, vol. 66, pp. 274–294, 2016.

[6] S. Akpinar, "Hybrid large neighbourhood search algorithm for capacitated vehicle routing problem," *Expert Systems with Applications*, vol. 61, pp. 28–38, 2016.

[7] G. B. Dantzig and J. Ramser, "The truck dispatching problem," *Management Science*, vol. 6, no. 1, pp. 80–91, 1959.

[8] J. J. Alcaraz, L. Caballero-Arnaldos, and J. Vales-Alonso, "Rich vehicle routing problem with last-mile outsourcing decisions," *Transportation Research Part E*, vol. 129, pp. 263–286, 2019.

[9] C. Lu, G. Kou, X. Zhou, Y. Peng, H. Sheng, and F. E. Alsaadi, "Time-dependent vehicle routing problem with time windows of city logistics with a congestion avoidance approach," *Knowledge-Based Systems*, vol. 188, pp. 1–8, 2020.

[10] P. Yan, H.-L. Luo, X. Li-ning et al., "A survey of vehicle routing optimization problems and solution methods," *Control Theory & Applications*, vol. 36, no. 10, pp. 1573–1584, 2019.

[11] International Energy Agency, *CO_2 Emissions from Fuel Combustion-Highlights*, International Energy Agency, Paris, France, 2015, https://www.iea.org/publications/freepublications/Publication/CO2EmissionsFromFuelCombustionHighlights2015pdf.

[12] G. Kou, X. Chao, Y. Peng, F. E. Alsaadi, and E. Herrera-Viedma, "Machine learning methods for systemic risk analysis in financial sectors," *Technological and Economic Development of Economy*, vol. 25, no. 5, pp. 716–742, 2019.

[13] C. Duan, C. Deng, A. Gharaei, J. Wu, and B. Wang, "Selective maintenance scheduling under stochastic maintenance quality with multiple maintenance actions," *International Journal of Production Research*, vol. 56, no. 23, pp. 7160–7178, 2018.

[14] X. Chao, G. Kou, Y. Peng, and F. E. Alsaadi, "Behavior monitoring methods for trade-based money laundering integrating macro and micro prudential regulation: a case from China," *Technological and Economic Development of Economy*, vol. 25, no. 6, pp. 1081–1096, 2019.

[15] S. A. H. Shekarabi, A. Gharaei, and M. Karimi, "Modelling and optimal lot-sizing of integrated multi-level multi-wholesaler supply chains under the shortage and limited warehouse space: generalised outer approximation," *International Journal of Systems Science: Operations & Logistics*, vol. 6, no. 3, pp. 237–257, 2018.

[16] S. A. H. Shekarabi, A. Gharaei, and M. Karimi, "Modelling and optimal lot-sizing of the replenishments in constrained, multi-product and bi-objective EPQ models with defective products: generalised cross decomposition," *International Journal of Systems Science: Operations & Logistics*, pp. 1–13, 2019.

[17] A. Gharaei, M. Karimi, and S. A. H. Shekarabi, "An integrated multi-product, multi-buyer supply chain under penalty, green, and quality control polices and a vendor managed inventory with consignment stock agreement: the outer approximation with equality relaxation and augmented penalty algorithm," *Applied Mathematical Modelling*, vol. 69, pp. 223–254, 2019.

[18] A. Gharaei, S. A. Hoseini Shekarabi, and M. Karimi, "Joint economic lot-sizing in multi-product multi-level integrated supply chains: generalized benders decomposition," *International Journal of Systems Science: Operations & Logistics*, pp. 1–13, 2019.

[19] N. Rincon-Garcia, B. Waterson, T. J. Cherrett, and F. Salazar-Arrieta, "A metaheuristic for the time-dependent vehicle routing problem considering driving hours regulations—an application in city logistics," *Transportation Research Part A: Policy and Practice*, vol. 137, pp. 429–446, 2020.

[20] M. Alinaghian and M. Naderipour, "A novel comprehensive macroscopic model for time-dependent vehicle routing problem with multi-alternative graph to reduce fuel consumption: a case study," *Computers & Industrial Engineering*, vol. 99, pp. 210–222, 2016.

[21] C. Malandraki, Time Dependent Vehicle Routing Problem: Formulations, Solution Algorithms and Computations Experiments Northwestern University, Evanston, IL, USA, 1989.

[22] R. Eglese, W. Maden, and A. Slater, "A road timetable to aid vehicle routing and scheduling," *Computers & Operations Research*, vol. 33, no. 12, pp. 3508–3519, 2006.

[23] A. V. Donati, R. Montemanni, N. Casagrande, A. E. Rizzoli, and L. M. Gambardella, "Time dependent vehicle routing problem with a multi ant colony system," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1174–1191, 2008.

[24] A. L. Kok, E. W. Hans, and J. M. J. Schutten, "Vehicle routing under time-dependent travel times: the impact of congestion avoidance," *Computers & Operations Research*, vol. 39, no. 5, pp. 910–918, 2012.

[25] Y. Huang, L. Zhao, T. Van Woensel, and J.-P. Gross, "Time-dependent vehicle routing problem with path flexibility," *Transportation Research Part B: Methodological*, vol. 95, pp. 169–195, 2017.

[26] Y. Kuo, "Using simulated annealing to minimize fuel consumption for the time-dependent vehicle routing problem," *Computers & Industrial Engineering*, vol. 59, no. 1, pp. 157–165, 2010.

[27] Z.-J. Ma, Y. Wu, and Y. Dai, "A combined order selection and time-dependent vehicle routing problem with time widows for perishable product delivery," *Computers & Industrial Engineering*, vol. 114, pp. 101–113, 2017.

[28] D. Taş, N. Dellaert, T. vanWoensel et al., "The time-dependent vehicle routing problem with soft time windows and stochastic travel times," *Transportation Research Part C: Emerging Technologies*, vol. 48, pp. 66–83, 2014.

[29] M. Andres Figliozzi, "The time dependent vehicle routing problem with time windows: benchmark problems, an efficient solution algorithm, and solution characteristics," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 3, pp. 616–636, 2012.

[30] X. Qu, J. Zhang, and S. Wang, "On the stochastic fundamental diagram for freeway traffic: model development, analytical properties, validation, and extensive applications," *Transportation Research Part B: Methodological*, vol. 104, pp. 256–271, 2017.

[31] Y. Bie, X. Xiong, Y. Yan, and X. Qu, "Dynamic headway control for high-frequency bus line based on speed guidance and intersection signal adjustment," *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 1, pp. 4–25, 2020.

[32] W. Qi, Y. Wang, Y. Bie, and J. Ren, "Prediction model for bus inter-stop travel time considering the impacts of signalized intersections," *Transportmetrica A: Transport Science*, pp. 1–19, 2020.

[33] Z. Liu, Y. Liu, Q. Meng, and Q. Cheng, "A tailored machine learning approach for urban transport network flow estimation," *Transportation Research Part C: Emerging Technologies*, vol. 108, pp. 130–150, 2019.

[34] X.-B. Hu, M.-K. Zhang, Q. Zhang, and J.-Q. Liao, "Co-evolutionary path optimization by ripple-spreading algorithm," *Transportation Research Part B: Methodological*, vol. 106, pp. 411–432, 2017.

[35] H. Wen, J. Wu, and Y. Duan, "A methodology of timing co-evolutionary path optimization for accident emergency rescue considering future environmental uncertainty," *IEEE Access*, vol. 7, pp. 131459–131472, 2019.

[36] F. Cheng-hong, Y. Shu-min, and Y. Zhang, "Promoted short-term traffic flow prediction model based on deep learning and support vector regression," *Journal of Transportation Systems Engineering and Information Technology*, vol. 19, no. 4, pp. 130–148, 2019.

[37] C. Ru-qing, L. Jia-chun, and Y. Jin-shou, "Short-term traffic flow forecasting based on hybrid FWADE-ELM," *Control and Decision*, pp. 1–8, 2019.

[38] J. Wang and Q. Shi, "Short-term traffic speed forecasting hybrid model based on chaos-wavelet analysis-support vector machine theory," *Transportation Research Part C: Emerging Technologies*, vol. 27, pp. 219–232, 2013.

[39] J. Y. T. Wang, M. Ehrgott, and A. Chen, "A bi-objective user equilibrium model of travel time reliability in a road network," *Transportation Research Part B: Methodological*, vol. 66, pp. 4–15, 2014.

[40] R. Faturechi and E. Miller-Hooks, "Travel time resilience of roadway networks under disaster," *Transportation Research Part B: Methodological*, vol. 70, pp. 47–64, 2014.

[41] X.-B. Hu and J.-Q. Liao, "Co-evolutionary path optimization by ripple-spreading algorithm," in *Proceedings of the 2016 IEEE Congress on Evolutionary Computation*, pp. 4535–4542, Canada, Vancouver, July 2016.

[42] X.-B. Hu, M. Wang, M. S. Leeson, E. A. Di Paolo, and H. Liu, "Deterministic agent-based path optimization by mimicking the spreading of ripples," *Evolutionary Computation*, vol. 24, no. 2, pp. 319–346, 2016.

[43] M. T. Goodrich, Algorithm Design: Foundations, Anal., and Internet Examples, Wiley, Hoboken, NJ, USA, 2001.

[44] M. Sniedovich, "Dijkstra's algorithm revisited: the dynamic programming connexion," *Journal of Control and Cybernetics*, vol. 35, pp. 599–620, 2006.

[45] M. Sniedovich, Dynamic Programming: Foundations and Principles, Francis & Taylor, Abingdon, UK, 2010.

[46] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, Prentice-Hall, Upper Saddle River, NJ, USA, 3rd edition, 2010.

[47] W. Qi, B. Shen, and L. Wang, "Model of driver's eye movement and ECG index under tunnel environment based on spatiotemporal data," *Journal of Advanced Transportation*, vol. 2020, Article ID 5215479, 11 pages, 2020.

[48] Z. Liu, S. Wang, K. Huang, J. Chen, and Y. Fu, "Practical taxi sharing schemes at large transport terminals," *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, pp. 596–616, 2019.

[49] Y. Liu, Z. Liu, and R. Jia, "DeepPF: a deep learning based architecture for metro passenger flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 18–34, 2019.

Research Article

Traffic Flow Modeling of Freeway Variable Speed Limit Control Based on the Big Data of Driving Behavior

Xu Qu [1,2] **Linheng Li,**[1,2] **Ziwei Yi,**[1,2] **Peipei Mao,**[1,2] **and Mofeng Yang** [3]

[1]*School of Transportation, Southeast University, Nanjing, China*
[2]*Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Road #2, Nanjing 211189, China*
[3]*Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742, USA*

Correspondence should be addressed to Xu Qu; quxu@seu.edu.cn

Variable speed limit (VSL) control is a flexible restriction on the rate at which motorists can drive on a given stretch of road. Effective VSL control can increase safety and provide clear guidance for motorists. Previous traffic flow models of VSL control were mostly based on the influence of VSL on average speed (macro) or driver's expected speed (micro). Few models considered the influence of VSL on driver's actual driving behavior. In this paper, we first briefly introduce the big traffic data involved in this study and explain the mapping relationship between the data and driving behavior. Then, we analyze the driver's actual driving behavior under the VSL control. Then, an improved single-lane cellular automaton model is established based on the driving behavior characteristics under VSL control. After that, we calibrate the parameters of the single-lane cellular automaton model with the left lane as the calibration object. Finally, this paper uses the proposed single-lane cellular automaton model to simulate the traffic flow characteristics under VSL control. The numerical simulation results show that the simulation of the variable speed limit in different density intervals presents different results, but these results are consistent with the actual situation of variable speed limit control, which verifies the validity of the proposed model.

1. Introduction

The variable speed limit (VSL) control system adjusts the road speed limit in real time based on the traffic flow state detected by advanced traffic flow and environment detection technology. As one of the advanced highway management control methods, VSL has been applied in Western countries such as Germany, the United States, Netherlands, the United Kingdom, Finland, Israel, New Zealand, and Australia [1]. The implementation effect indicates that VSL has a great impact on improving road traffic safety and alleviating traffic congestion [2–5].

The essential basic research content of VSL is traffic flow modeling. Macroscopic models are an important component of VSL modeling, which can be divided into two categories in terms of the characteristics of different VSL traffic flows [6–9]. One category is well studied, considering the influence of VSL on the equilibrium speed relationship [10–14]. The equilibrium speed function is replaced by a new one with the VSL control parameter according to the modified fundamental diagram under VSL control. Another category establishes the macroscopic model using the correspondence between macroscopic and microscopic parameters when simulating the driver behavior under VSL control.

The macroscopic model has the advantage of high operational efficiency in numerical simulation, but most of the existing macroscopic models lack the behavior description of individual vehicles, which makes it impossible to simulate the safety benefits of VSL. Hence, the optimization goal of VSL is limited to efficiency.

The microscopic model is preferred because both the efficiency benefits and security benefits of VSL strategies can be evaluated. Lee et al. [15] used PARAMICS to simulate

VSL control; Torday and Bierlaire [16] modeled the VSL microscopic model based on the MITSIM simulator; and Park and Yadlapati [17] used VISSIM to model the VSL with different compliance rates. Other microscopic traffic simulations of Cruise Control (CC) and Adaptive Control (ACC) are used [18–20]; such models were used for freeway traffic for the evaluation of the effectiveness of VSL [21, 22]. However, in most of the microscopic models, the phenomenon that VSL affects the vehicle's desired speed is pondered, but the changes in drivers' behavior are ignored. In fact, VSL performance depended greatly on compliance rates [23]; the VSL not only changes the desired speed but also changes the following behavior of drivers. This is reflected in a previous paper where the proportion of the smaller headway was reduced [24]. Besides, models on considering the desired speed are not reasonable. In most of the studies, vehicles' speed was set equal to the VSL value, which does not correspond with the real estimation. When simulating large or small value of VSL control, the desired speed would be slightly above or beneath the limit value, which is worth exploring but had not been investigated in existing models. Therefore, it is necessary to explore the impact regularity of VSL control under different traffic conditions and compliance rates.

Traffic data always plays an important role in the process of traffic flow modeling; especially in the research of microscopic traffic flow modeling, it is necessary to describe driving behavior characteristics through massive microscopic traffic data. Recently, many scholars have tried to use driving behavior data as a research entry point to solve a series of related transportation problems. Li et al. [25] analyzed drivers' deceleration behavior based on naturalistic driving data and, finally, a brake control system was designed. In order to evaluate whether a driving behavior is fuel-efficient, Hao et al. [26] mined the traffic operating data and fuel consumption data and finally built a corresponding analysis model. Deng et al. [27] established a curve safe speed model on the basis of a thorough analysis of driving behavior data.

The use of empirical data to investigate driving behaviors with VSL control was well-studied [10, 12, 17]. The consensus emerging from these studies is that VSL systems can improve safety by decreasing the mean speed value, the speed difference, and the percentage of small-time headway and can increase the mean time headway. Studies using empirical traffic data as input can objectively analyze the impacts of VSL on driving behaviors and derive credible conclusions. However, with the constraint of predefined VSL control strategies, the majority of these studies only have access to data from a limited range of traffic conditions and limited VSL values. Therefore, most of the researchers merely performed qualitative comparisons between driving behavior parameters whether they used VSL control or not while missing the quantitative examination of the relationship between traffic variables and VSL values. In a real-world implementation, VSL values vary in accordance with the prevailing traffic conditions and weather conditions. However, few scholars evaluated the potential outcomes of each VSL value. In the meantime, only a limited number of

dissertations controlled traffic states when examining VSL impacts; therefore, it is difficult to determine whether the obtained results are caused by VSL control or by the difference in traffic states within the collected empirical data.

This paper summarized the literature on traffic flow models and driving behavior under VSL control and proposed an improved cellular automaton model that represented individual driving behavior and reproduced corresponding traffic flow characteristics under VSL control. The rest of this paper is organized as follows. Driving behavior under VSL control is analyzed with data from a motorway in Section 2. Section 3 establishes the single cellular automaton model taking account of the compliance parameter. In Section 4, empirical data of a highway segment is used to identify the distribution of driver's individual radical degree and calibrate the parameters of the proposed model. In Section 5, numerical simulation is conducted to validate the superiority of the proposed model in reproducing traffic flow characteristics under VSL control.

2. Big Data Analysis of Driving Behavior

2.1. Data Preparation. The data used in this study was collected from a two-way European motorway segment with three lanes in each direction. The traffic flow state changes obviously, including recurrent and nonrecurrent congestions. The traffic flow information, such as the vehicular speed, the headway, and the vehicle length, was collected by loop detectors installed every 500–600 m in each lane. In order to alleviate congestion and improve driving safety, the expressway was equipped with VSL control facilities that adopt a control strategy based on mitigating traffic shocks. A wide variety of speed limits including 50 km/h, 60 km/h, 80 km/h, 100 km/h, and 120 km/h were implemented in the system.

This research collected traffic flow data and corresponding VSL control speed data from 7:00 AM to 6:00 PM for two weeks at a nonweaving area site. Only data in good weather condition (no rain or fog) and visibility conditions (at daytime) was selected to avoid external interference. Missing and wrong data was removed. Finally, the dataset including 4,266 minutes of data samples with 355,599 vehicles was constituted.

The traffic state is a crucial factor when investigating the VSL impacts on driving behavior. Under various traffic states, driving behavior will be differently affected by the VSL system. Hence, it is difficult to determine whether the obtained results are caused by the VSL control or by the difference in traffic states through the raw empirical data. To eliminate the impact of traffic states, the classification of the samples according to traffic states is indeed critical. In this paper, the collected samples were firstly classified into different traffic density intervals to measure traffic congestion. Then the traffic data in the same density interval with different VSL control was compared to that without VSL control to investigate the effect of VSL control on driving behavior.

2.2. Desired Speed Distribution. From the previous study, it can be inferred that the VSL has the ability to reduce the desired speed, since the average speed in each density interval under free-flow state was descended with different VSL values. Gartner's study mentioned that when the space headway is greater than 125 m (corresponding to the headway of 4~6 s at 20~30 m/s vehicle speed), the vehicle will not be in the following state [28]. According to the conclusion, vehicle speed in the above-mentioned state reflects the driver's desired speed, which is not affected by the preceding vehicle. A statistical analysis of the vehicle speed with a headway greater than 5 s in the free-flow state was performed. Figure 1 shows the vehicle speed distribution of the left lane in the uncontrolled state and VSL control state including 120 km/h, 100 km/h, and 80 km/h.

Figure 1 demonstrates that the effect of the VSL on the driver's desired speed is holistic. After setting VSL, the driver basically responded and reduced the speed. The exponential regression line in the figure shows that the driving speed gradually increases as the headway increases, which is in line with reality. At the same time, it can be seen from the figure that under different VSL values, a few drivers drive at a speed obviously higher than the VSL value. It is believed that drivers who are significantly speeding are relatively aggressive drivers who may not follow VSL. Besides, at the high-speed limit, these drivers will be speeding by more than 30%, while the percentage will rise to 50% at the low VSL value. We believe that some drivers might challenge the rationality of VSL control when a low-speed limit is implemented, whereby the reduced VSL obedience rate leads to the phenomenon of higher mean speeds under lower VSL values.

2.3. Average Desired Speed. The average desired speed reflects the compliance rate of drivers to some extent. In order to quantify the change in the average expected speed, the arithmetic average speed is calculated in the case of different speed limit values on different lanes.

Table 1 shows the average desired speed with the headway time greater than 5 s under different VSL values in each lane ("—" indicates that the number of samples is less than 50; no statistics were performed). It can be observed that the driver's desired speed decreased significantly with VSL in all lanes compared with the uncontrolled state. In addition, from the speed limit value of 120 km/h to 60 km/h, the average speed decreases as the variable speed limit decreases. However, the reduction is not linear, since the difference between the driver's desired speed and the VSL value is bigger under the control of the lower VSL value like 60 km/h.

2.4. Small Headway Distribution. The distribution of the headway, especially the proportion of the smaller headway, is critical to the stability of the traffic flow. This paper mainly focuses on the effect of the variable speed limit on the percentage of vehicles whose headway is less than 1 second. In order to quantify the effects of different VSL values on the headway, we calculated the percentage of small headways within different density intervals under different speed limits on different lanes. Table 2 summarizes the percentage of vehicles with smaller than 1 s headway under all combinations of density intervals and speed limits on the left, the middle, and the right lane ("—" indicates that less than 300 vehicles were observed within the associated density interval).

It can be observed from the table that the ratio of less than 1 s headway in the uncontrolled state increases with the increase of the density, while the ratio under the VSL control does not change significantly with the increase of the density. In general, the percentage of less than 1 s headway in the VSL control is obviously lower than that in the uncontrolled state. The percentage at the speed limit of 100 km/h and 80 km/h is the least, not even less than 50% of the uncontrolled state. The results show that some drivers become more cautious under VSL control and tend to maintain a large headway to ensure safety.

3. Single-Lane Cellular Automaton Model

This section presents a single-lane cellular automaton model for variable speed limit (VSL) control. On the basis of improved cellular automaton model incorporating impaired driver's radical feature (RF) raised by Qu et al. [29], this paper introduces the VSL compliance parameter to describe the driver's compliance characteristics, setting the desired speed according to the compliance rate of each vehicle, and then modify the moderation probability of the RF model combined with compliance rate; that is, if a driver with VSLs compliance and within the range of the VSLs control keeps headway less than 1 second, the driver will possess higher decelerating probability.

In the model, the driver will possess a higher decelerating probability while being within the range of the VSL control with headway less than 1 second. The model was established under two assumptions:

(1) The VSL control changes the desired speed of most drivers, but the desired speed and the value of VSL are not equal and their relationship is not linear.

(2) The ratio of the smaller headway (less than 1 s) is significantly reduced under the VSL control [24], which reveals that some drivers become more cautious in driving and tend to maintain a larger headway to ensure safety.

The random probability of deceleration expression is

(a)

(b)

(c)

(d)
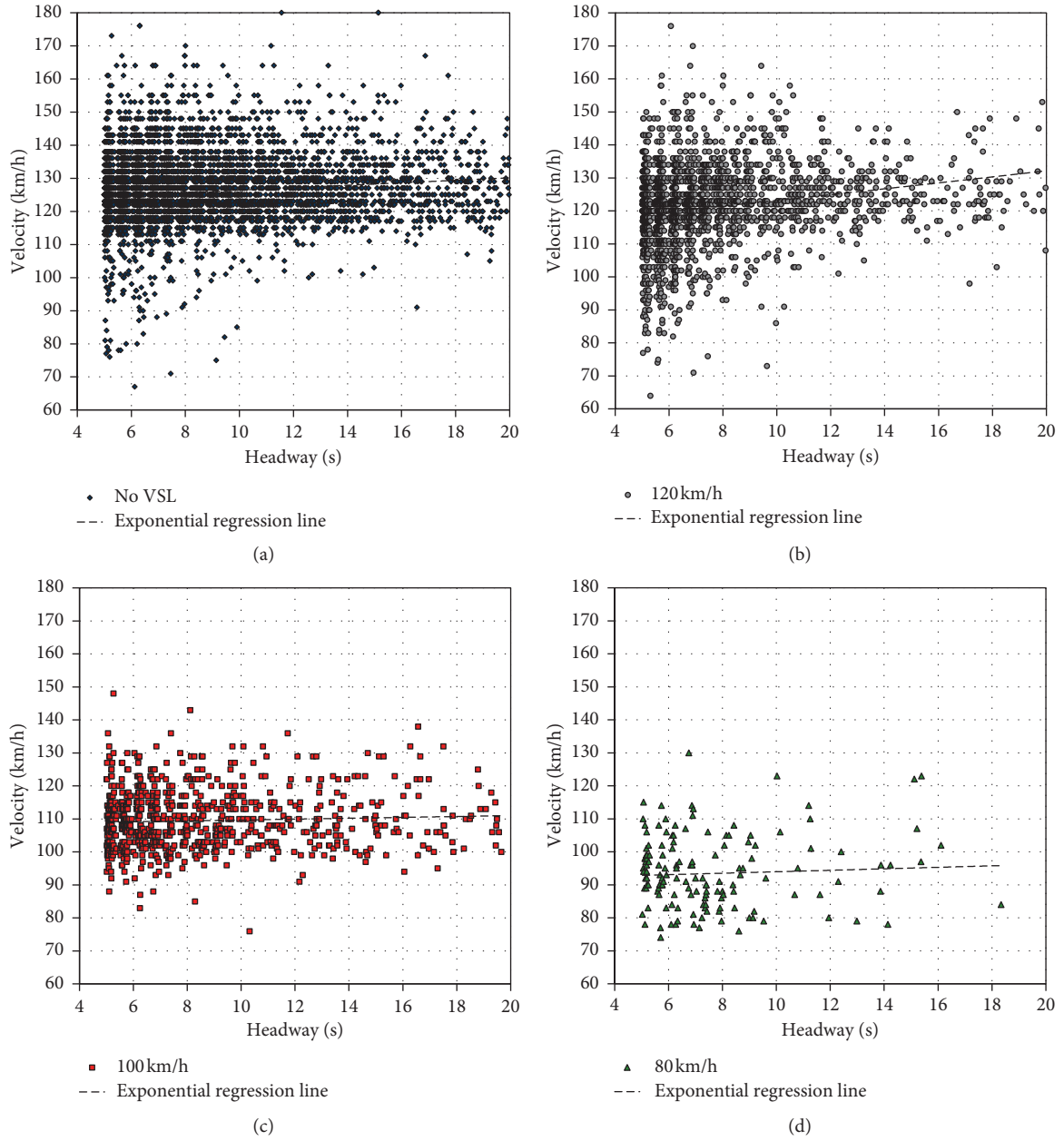
Figure 1: Speed headway of the vehicles on the left lane (headway greater than 5 s): (a) uncontrolled state, (b) 120 km/h, (c) 100 km/h, and (d) 80 km/h.

$$P_{\text{slow}}\left(v_n(t), b_{n+1}(t), t_h, t_s, c_n, a_n(t)\right) = \begin{cases} p_b: \text{if } b_{n+1} = 1 \text{ and } t_h < t_s \\ p_0: \text{if } v_n = 0 \\ p_c: \text{if } c_n = 1 \text{ and } a_n(t) = 1 \text{ and } t_h < 1 \\ p_d: \text{in all other cases,} \end{cases} \tag{1}$$

where c_n is the compliance parameter of nth vehicle, $c_n = 1$ means compliance, $c_n = 0$ indicates noncompliance, the compliance of vehicle is affected by the driver's personal characteristics, and the compliance property of each vehicle is initially determined and will not change along with the simulation or varies after each update circle. The compliance

parameter of each vehicle is determined by the following formula:

$$c_n = \begin{cases} 1: \text{if } r \text{ and } (0, 1) < C \\ 0: \text{else.} \end{cases} \tag{2}$$

TABLE 1: The average vehicle speed with a headway greater than 5 s at different VSLs in each lane.

Lane	Average desired speed (km/h)					
	No VSL	120 km/h	100 km/h	80 km/h	60 km/h	50 km/h
Left lane	126.9	121.8	109.8	94.0	86.1	—
Middle lane	114.9	105.0	99.0	85.1	80.2	—
Right lane	97.0	89.9	88.3	78.8	74.4	—

TABLE 2: Percentages of smaller than 1 s headway within different density intervals under different speed limits.

Lane	Density intervals (veh/km)	Percentage of less than 1 s headway under different speed limits					
		No VSL	120 km/h	100 km/h	80 km/h	60 km/h	50 km/h
Left lane	0–5	24%	21%	15%	—	—	—
	5–10	32%	28%	19%	15%	—	—
	10–15	35%	26%	14%	10%	—	—
	15–20	37%	26%	14%	11%	15%	20%
	20–25	36%	29%	14%	16%	18%	—
Middle lane	25–30	9%	6%	—	—	—	—
	5–10	15%	11%	7%	—	—	—
	10–15	22%	16%	12%	7%	—	—
	15–20	26%	20%	14%	12%	14%	14%
	20–25	23%	22%	14%	—	—	—
Right lane	25–30	3%	3%	—	—	—	—
	5–10	7%	6%	4%	3%	—	—
	10–15	13%	11%	5%	4%	5%	—
	15–20	18%	14%	—	—	—	—

C is the VSL compliance rate, defined as the percentage of the drivers affected by VSL control, abbreviated as "compliance rate" afterward. "r and $(0,1)$" indicates a random number between 0~1. $a_n(t)$ is a parameter to evaluate whether the vehicle is within the VSL control range, $a_n(t) = 1$ indicates that the nth vehicle is within the VSL control range at time t, while $a_n(t) = 0$ indicates that it is out of the range. Other parameters are the same as the RF model defined.

The update rules of the proposed model are as follows:

(1) Determining the desired speed as

if $\left(c_n = 1 \text{ and } Xvsl_i \leq x_n(t) < Xvsl_{i+1}\right)$,

then: $v_{\text{des}}(n) = \min\left(\bar{v} + \alpha_n, v_{\max,n}\right)$ else $v_{\text{des}}(n) = v_{\max,n}$.

$$(3)$$

(2) Determining the probability of randomization as

$$P_{\text{slow}}\left(v_n(t), b_{n+1}(t), t_h, t_s, c_n, a_n(t)\right),$$
$$b_n(t+1) = 0,$$
$$a_n(t+1) = 0.$$
$$(4)$$

(3) Acceleration is

$$v_n(t+1) \longrightarrow \min\left(v_n(t) + 1, v_{\text{des}}(n)\right).$$
$$(5)$$

(4) Deceleration is

$$v_n(t+1) \longrightarrow \min\left(v_n(t), d_n^{(\text{eff})}\right),$$
$$\text{if } \left(v_n(t+1) < v_n(t)\right) \quad \text{then: } b_n(t+1) = 1.$$
$$(6)$$

(5) Randomization is

$$\text{if } \left(\text{rand} < p_{\text{slow}}\right),$$
$$\text{then: } v_n(t+1) \longrightarrow \max\left(v_n(t+1) - 1, 0\right).$$
$$(7)$$

(6) Location update is

$$x_n(t+1) \longrightarrow x_n(t) + v_n(t+1).$$
$$(8)$$

(7) Determining whether it is in the range of VSL control as

$$\text{if } \left(Xvsl_i \leq x_n(t+1) < Xvsl_{i+1} \quad \text{and} \quad Vvsl_i \neq 0\right),$$
$$\text{then: } a_n(t+1) = 1,$$
$$(9)$$

where $Xvsl_i$ is the starting position of the ith VSL control area. $\bar{v}$ denotes the average value of vehicle's desired speed (AVDS) under the VSL control, determined by the speed limit value in the ith VSL control area, which is called the AVDS of VSL, described as $\bar{v} = f(Vvsli)$. $Vvsl_i$ indicates the

speed limit value in the ith VSL control area. $v_{\text{des}}(n)$ is the desired speed of the nth vehicle. $v_{\text{max},n}$ is the maximum speed of the nth vehicle. The other parameter definitions are the same as the comfortable driving (CD) model [30].

4. Model Calibration

4.1. AVDS and Average Maximum Speed (AMS) under VSL. In this study, the left lane was selected as the calibration target lane for the single-lane model. The vehicle's desired speed and the VSL values are neither equal nor linear under the VSL control. Therefore, the AVDS under the different VSL control cases needs to be calibrated according to the actual data. On the basis of a number of samples, the average speed of the vehicle with headway larger than 5 seconds can be selected as the value of AVDS under a certain speed limit control. If it is not possible to obtain a sufficient sample size, the average speed in the lower density state may be selected as the value of AVDS. In this study, the AMS is equivalent to the AVDS under the condition of no VSL control, and the calculation method is the same as in calculating AVDS.

For the speed limit values of 120 km/h, 100 km/h, 80 km/h, and 60 km/h, we selected the average speed of the vehicle with headway greater than 5 seconds as the AVDS of the VSL control. For the 50 km/h speed limit value, the average speed in the lower density state was selected as its AVDS. The specific values are shown in Table 3.

The result of analyzing exiting data reveals that the formation of regression function is consistent with quadratic polynomial; in order to supplement the missing values to calculate the corresponding AVDS, the quadratic polynomial regression can be performed through the existing data, and the values are listed in Table 3. Herein, only the speed limit values, which can be integers divided by 10 and greater than 50 km/h, are listed. The regression function is modeled as (10), where x is cell speed value and y is the corresponding AVDS. The coefficient of determination is 0.97:

$$y = 0.0010x^2 - 1.213x + 126.3. \tag{10}$$

4.2. Value of Compliance Rate. The value of the compliance rate is affected by many factors, among which the VSLs control traffic regulations and the driver's recognition are the most significant. In general, the compliance rate under the mandatory regulations is superior to recommended regulations and is more prominent along with the penalty rate rising or driver's recognition improvement.

There are some drawbacks of analyzing compliance through a survey. Firstly, it depends on whether the driver answers the questionnaire honestly. What is more, driving behavior may be different between humans. Therefore, in this paper, the statistics method was conducted. Through observing the distribution of vehicle speeds with its headway larger than 5 seconds, we found that some vehicles' speeds are significantly higher than others in any speed limit control. These vehicles have noncompliance characteristics because they were not affected by VSL control cases.

It can be concluded that, among all vehicles with headway larger than 5 seconds, except for some vehicles with speed beyond the critical value, the percentage of the remaining vehicles is the compliance rate of VSL control. The critical speed value should be kept in a reasonable scope to decrease the errors in statistical analysis; in this study, the average speed over 30% of the limit value was selected. In the case of a 100 km/h limit control situation, according to statistics results, 5% of vehicles were beyond 130 km/h. As a consequence, the compliance rate is 95%, drafted as $C = 95\%$.

4.3. Calibration of Random Probability of Deceleration p_c. The random probability of deceleration p_c is to simulate the driver's cautious driving characteristics under the VSL control; its value will directly affect the proportion of the small headway. Therefore, we will calibrate the parameter p_c by determining the ratio of the headway less than 1 second in the actual VSL control.

To calibrate the p_c values corresponding to different speed limit values, we performed simulations under speed limit values of 120 km/h, 100 km/h, 80 km/h, 60 km/h, and 50 km/h, respectively. For a certain speed limit value, the value of the p_c is gradually changed, and other parameters were fixed. The road length was set to 4000 cell units, the time step was set to 1 second, and the density was set to 15 veh/km, respectively. Periodic boundary conditions were adopted. The initial position and speed of the vehicle were randomly generated. Each simulation runs 10,600 time steps, where the first 10,000 time steps were not counted in order to eliminate the effects of transients, while in the next 600 time steps, the ratio of the headway less than 1 second was calculated. For the same p_c value, the simulation was taken 10 times to obtain an average value. Figure 2 shows the ratio of the headway less than 1 second at different p_c values under the 100 km/h speed limit control.

It can be seen from the figure that the ratio of the headway less than 1 second is relatively sensitive to the parameter p_c. Qu et al. [24] mentioned that, in practice, the ratio of the headway less than 1 second under the control of 100 km/h speed limit in the free-flow state is about 15%~20%; therefore, the value of p_c ranges from approximately 0.5 to 0.8 under the limit speed value of 100 km/h. In the same way, we calibrate the p_c values of other speed limit values. The results are shown in Table 4.

As can be seen from the results, the p_c values corresponding to the excessively high- and low-speed limit values are small, while under the moderate speed limit, values (80 km/h, 100 km/h) are large. This reveals that the driver's cautious driving behavior under the control of the moderate speed limit is the most obvious. For the sake of simplicity, this study divides the value of the parameter p_c into two parts: when the speed limit value is between 80 km/h and 100 km/h, $p_c = 0.6$; under other speed limit values, $p_c = 0.4$.

5. Numerical Simulation of the Model

In this section, the simulation of a fundamental diagram, time-space diagram, and speed differences characteristics was conducted to evaluate the performance of the single-lane cellular automaton model. In the simulation, the length of the road was set to 4000 cell units, and the time step was

TABLE 3: Variable speed limit AVDS values of the left lane.

Speed limit value (km/h)	AVDS (km/h)	Corresponding cell speed value (cell/s)
No VSL	126.9	23
120	121.8	22
110	113.8	21
100	109.8	20
90	98.1	18
80	94.0	17
70	90.3	17
60	86.1	16
50	90.6	17
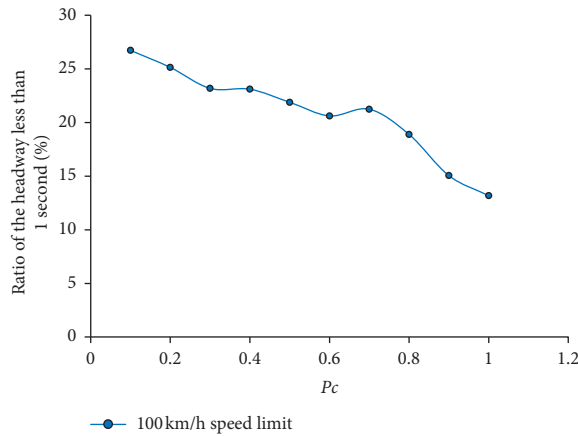
FIGURE 2: Proportion of the headway less than 1 second at different p_c values.

TABLE 4: Range of values of different VSL parameters, p_c.

Speed limit value (km/h)	The range of p_c
120	0.3–0.5
100	0.5–0.9
80	0.5–0.8
60	0.3–0.6
50	0.3–0.5

set to 1 second with periodic boundary condition. The initial position and speed of the vehicle were randomly generated. Each simulation ran 10,600 time steps, of which the first 10,000 time steps did no statistics to eliminate the impact of the transient. The value for the new model's parameters is the same as that in the calibration session.

5.1. Fundamental Diagram. This study carried out a pseudo-numerical simulation of VSL control with values of 120 km/h, 100 km/h, and 80 km/h with no VSL control, which has been run 10 times for each density value; then the average traffic flow was calculated. The fundamental diagram under different speed limit values is shown in Figure 3. It can be seen from the figure that, in the low-density region, the average speed (curve slope) under the VSL control is minor compared to that of the uncontrolled situation, with the tendency of the speed limit value and the average speed being the same. In the high-density region, it can be found
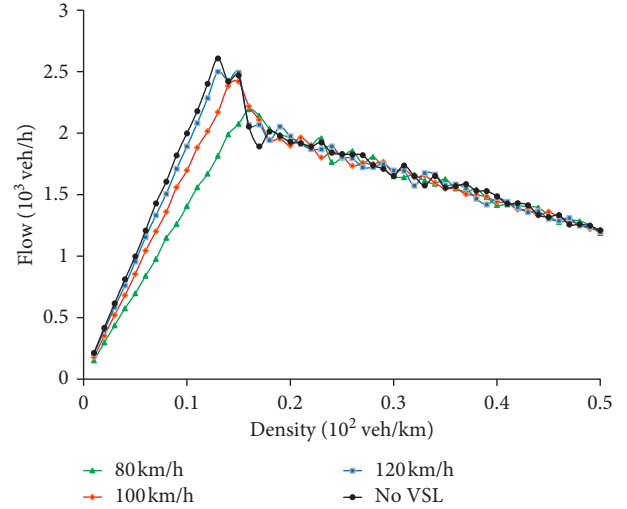

FIGURE 3: Fundamental diagram of different speed limit values for single-lane cellular automaton model simulation.

that the fundamental diagram at different speed limit values does not significantly differ from the uncontrolled situation. In the moderate density region, the critical density (the density region where the traffic flow changes phase) under the VSL control is larger than that in an uncontrolled situation, and the corresponding key density value is growing with the descent of the speed limit value. Accompanied by the decrease of speed limit value (80 km/h), the characteristics of the traffic flow phase transition have become vague. In addition, under the VSL control, the road capacity declined with its scope being weightier as the speed limit value decreased. The features of the fundamental diagram in the above simulations are basically consistent with those in actual observation.

5.2. Time-Space Diagram. Figure 4 shows the time-space diagram of the 100 km/h speed limit and no VSL control cases at 0.1, 0.15, and 0.3 density, respectively. It can be seen from the figure that, in the small density ($k = 0.1$) and large density ($k = 0.3$) regions, the VSL control does not significantly change the basic form of the time-space diagram. In the small density region, the traffic flow is in the free-flow state as a form of the vehicle fleet, while the large density region exhibits a wide motion-blocking feature with a large-scale blockage group. In the moderate density ($k = 0.15$) region, the VSL control significantly changes the shape of the time-space diagram, because it does not exhibit a phase separation phenomenon consisting of a distinct blocking phase and a moving phase compared to the uncontrolled situation, but possesses free-flow characteristics. It can be confirmed in Figure 4 that when $k = 0.15$, the free-flow state changes to the blocking phase without VSL control and maintains stability while being under control. This corresponds with a previous study reporting that VSL can increase the critical density [12].

It can be found that, for a certain speed limit value, the spatiotemporal characteristics changed every time, but some

(a)

(b)

(c)

(d)

(e)

(f)
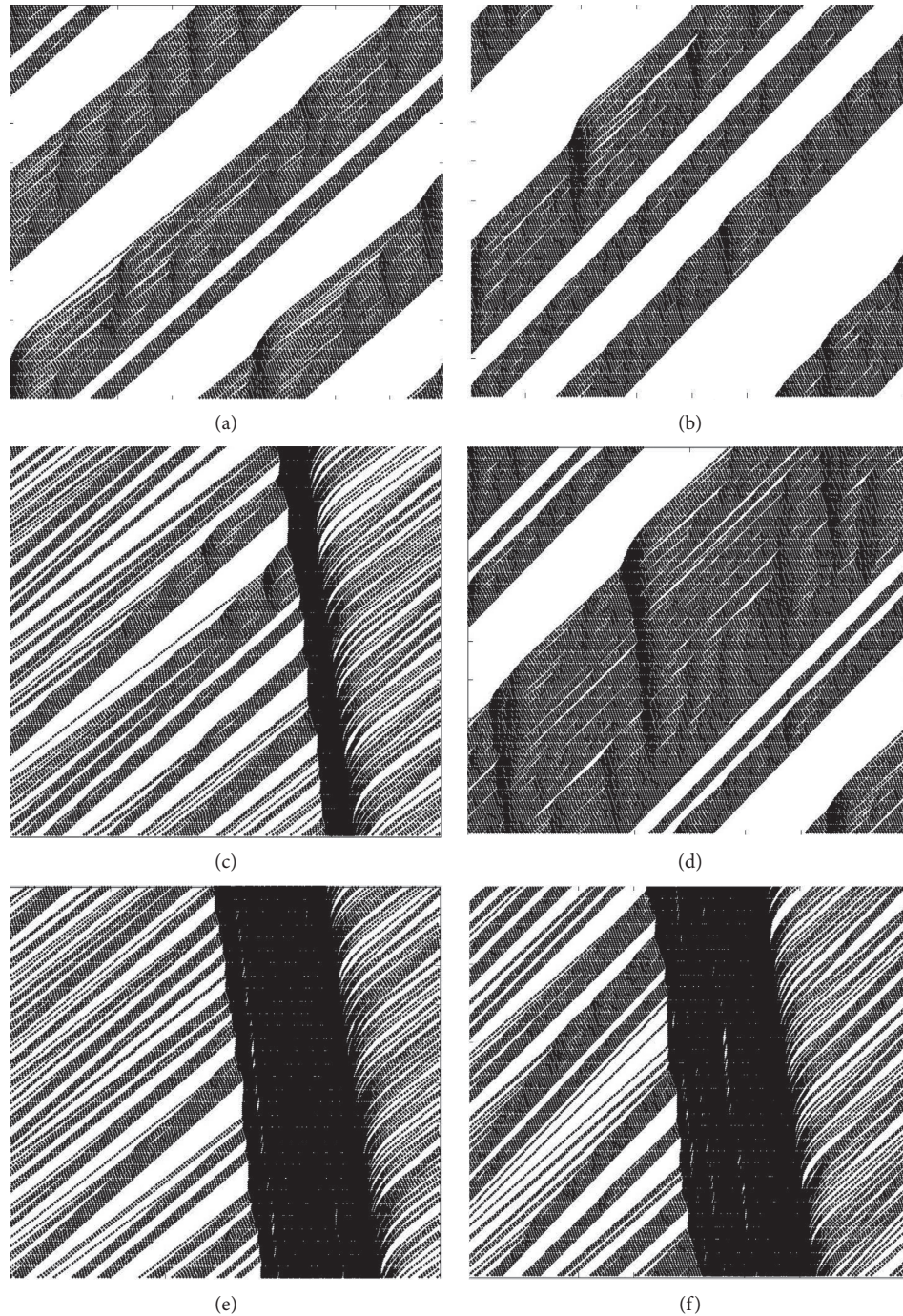
FIGURE 4: Time-space diagrams for controlled and uncontrolled situations. (a, c, e) Time-space diagrams at $k = 0.1$, 0.15, and 0.3 density in uncontrolled state. (b, d, e) Time-space diagrams of VSL = 100 km/h at $k = 0.1$, 0.3, and 0.5 density.

regulations existed. To explore the discipline, we carried out simulations to calculate the frequency of the free-flow state. On different density values near the critical density, 20 times of simulations under each speed limit value were conducted, the results are shown in Table 5.

The result reveals that the speed limit value decreased as the density increased, and the probability of generating a free-flow state expanded. This indicates that the VSL control

could boost the critical density of traffic flow, which is consistent with the fundamental diagram and the actual observation of traffic flow characteristics.

5.3. Average Speed Difference of Two Adjacent Vehicles. The average speed difference (ASD) of two adjacent vehicles was selected to measure the speed difference characteristics of traffic flow. It is defined as the mean value of speed

Table 5: Proportion of free-flow characteristics at different density limits in time-space diagrams.

Density value (veh/km)	No VSL (%)	120 km/h (%)	100 km/h (%)	80 km/h (%)	60 km/h (%)
0.14	100	100	100	100	100
0.15	25	30	85	100	100
0.16	10	10	35	100	100
0.17	0	5	10	95	100
0.18	0	0	0	20	90
0.19	0	0	0	0	25
0.20	0	0	0	0	0

Figure 5: ASD of two adjacent vehicles in different speed limit values and different density conditions.

difference between the adjacent vehicles passing a given point at a specified time interval [31], as shown as follows:

$$\text{ASD} = \overline{\Delta v} = \frac{\sum_{n=1}^{N-1} |v_n - v_{n+1}|}{n-1}, \tag{11}$$

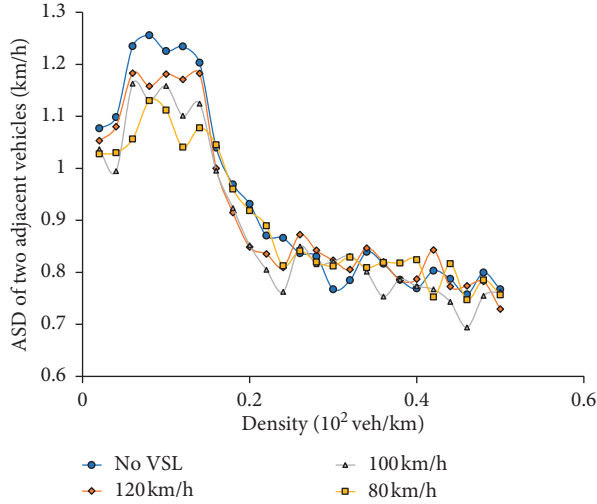where v_n represents the speed of the nth vehicle when passing the fixed point and N represents the number of vehicles passing a given point. In this study, simulations were carried out 10 times for each density condition following the single-lane cellular automaton model. Through simulations, the ASD of two adjacent vehicles in different speed limit values and different density conditions is shown in Figure 5.

It can be seen from the figure that the ASD of two adjacent vehicles in different VSL control cases is basically similar to the trend of the density varying. Except for excessive small density conditions, the ASD of two adjacent vehicles decreased with the density growth. In the small density region, the value under the VSL control is less than that in an uncontrolled state; this tendency is similar to actual situations [24]. In the large density region, different speed limit values are not significantly different.

6. Conclusions

This paper studies the real driving behavior data under VSL control and finds the following conclusions:

(1) Through the processing and analysis of traffic big data, the traffic data in the same density interval with different VSL control was compared with that without VSL control to investigate the effect of VSL control on driving behavior

(2) Under the control of the VSL, there always exists a certain proportion of drivers who fail to comply with the VSL control

(3) The VSL control changes the driver's desired speed, but the average desire speed and the value of VSL are not equal and show a nonlinear relationship

(4) The proportion of smaller headway time (less than 1 s) under VSL control is significantly reduced, which shows that some drivers become more cautious under the VSL control

Based on the above-mentioned driving behavior, this study introduces the compliance rate parameter during the modeling process. In addition, it improves the slowing probability of the RF model [26], so that the vehicle slows down with a higher probability when the headway time is less than 1 s, in order to simulate the prudent driving behavior of the vehicle under the VSL control. In this study, the parameters of the model are calibrated using real vehicle data under VSL control. The numerical simulation results show that the model can not only fit the macroscopic traffic phenomena such as the fundamental diagrams and the time-space diagrams under the VSL control, but also fit the influence of the VSL on the speed dispersion.

It should be noted that this study assumed that the driver's radical feature is subject to the normal distribution when defining the function of radical feature parameter. Whether this assumption is reasonable or not needs to be discussed further in future research. Besides, this study is just early stage work to incorporate impaired driving behavior. More work should be done to develop a more accurate calibration process as well as multiple lane models. In our future research, we will focus on collecting real data to help discuss the driver's radical degree and its distribution under the car-following state.

Data Availability

The basic data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

Acknowledgments

References

[1] C. Levecq, B. Kuhn, and D. Jasek, "General guidelines for active traffic management deployment," Interim Report, Texas A&M Transportation Institute, College Station, TX, USA, 2011.

[2] Y. Zhang and P. A. Ioannou, "Combined variable speed limit and lane change control for highway traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 7, pp. 1812–1823, 2017.

[3] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Local feedback-based mainstream traffic flow control on motorways using variable speed limits," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1261–1276, 2011.

[4] M. Yu and W. "David" Fan, "optimal variable speed limit control in connected autonomous vehicle environment for relieving freeway congestion," *Journal of Transportation Engineering, Part A: Systems*, vol. 145, no. 4, Article ID 04019007, 2019.

[5] H.-Y. Jin and W.-L. Jin, "Control of a lane-drop bottleneck through variable speed limits," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 568–584, 2015.

[6] M. Hajiahmadi, G. S. van de Weg, C. M. J. Tampere et al., "Integrated predictive control of freeway networks using the extended link transmission model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 65–78, 2016.

[7] A. Csikós and B. Kulcsár, "Variable speed limit design based on mode dependent Cell Transmission Model," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 429–450, 2017.

[8] M. Kontorinaki, I. Karafyllis, and M. Papageorgiou, "Global exponential stabilisation of acyclic traffic networks," *International Journal of Control*, vol. 92, no. 3, pp. 564–584, 2019.

[9] J. R. José, I. Papamichail, M. Papageorgiou, and B. De Schutter, "Macroscopic modeling of variable speed limits on freeways," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 15–33, 2019.

[10] S. Smulders, "Control by variable speed signs: the Dutch experiment," in *Proceedings of the 6th International Conference on Road Traffic Monitoring and Control*, London, UK, 1992.

[11] M. Cremer, "Der verkehrsfluss auf schnellstraßen," in *Modelle, Überwachung, Regelung*, Springer-Verlag, Berlin, Germany, vol. 3.Theoretical Computer Science, 1979.

[12] M. Papageorgiou, E. Kosmatopoulos, and I. Papamichail, "Effects of variable speed limits on motorway traffic flow," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2047, no. 1, pp. 37–48, 2008.

[13] A. Hegyi, B. De Schutter, and H. Hellendoorn, "Model predictive control for optimal coordination of ramp metering and variable speed control ∗," in *Proceedings of the 1st European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems (EUNITE 2001)*, pp. 1–6, Tenerife, Spain, December 2001.

[14] I. Papamichail, K. Kampitaki, M. Papageorgiou, and A. Messmer, "Integrated ramp metering and variable speed limit control of motorway traffic flow," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 14084–14089, 2008.

[15] C. Lee, B. Hellinga, and F. Saccomanno, "Evaluation of variable speed limits to improve traffic safety," *Transportation Research Part C: Emerging Technologies*, vol. 14, no. 3, pp. 213–228, 2006.

[16] A. Torday and M. Bierlaire, "Papabiles: simulation-based evaluation of the impact of telematics in the lausanne area: a pilot study," in *Proceedings of the Swiss Transport Research Conference*, Ascona, Switzerland, March 2001.

[17] S. Yadlapati and B. Park, "Development and testing of variable speed limit control logics for work zones using simulation," Center for Transportation Studies at the University of Virginia, Charlottesville, VA, USA, 2004.

[18] V. Milanés and S. E. Shladover, "Handling cut-in vehicles in strings of cooperative adaptive cruise control vehicles," *Journal of Intelligent Transportation Systems*, vol. 20, no. 2, pp. 178–191, 2016.

[19] V. Milanes, S. E. Shladover, J. Spring, C. Nowakowski, H. Kawazoe, and M. Nakamura, "Cooperative adaptive cruise control in real traffic situations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 296–305, 2014.

[20] V. Milanés and S. E. Shladover, "Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data," *Transportation Research Part C: Emerging Technologies*, vol. 48, pp. 285–300, 2014.

[21] X. Y. Lu and S. Shladover, "MPC-based variable speed limit and its impact on traffic with V2I type ACC," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 3923–3928, Maui, HI, USA, 2018 November.

[22] M. Morari and J. H. Lee, "Model predity control: past, present and future," *Neuroscience*, vol. 23, pp. 667–682, 2014.

[23] J. Ma, X. Li, S. Shladover et al., "Freeway speed harmonization," *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 1, pp. 78–89, 2016.

[24] X. Qu, W. Wang, B. Ran, and Y. Dai, "Quantitative analysis of the impact of variable speed limits on motorway safety," in *Proceedings of the Transportation Research Board 96th Annual Meeting*, Washington DC, USA, 2017.

[25] S. Li, P. Li, Y. Yao, X. Han, Y. Xu, and L. Chen, "Analysis of drivers' deceleration behavior based on naturalistic driving data," *Traffic Injury Prevention*, vol. 21, no. 1, pp. 42–47, 2020.

[26] R. Hao, H. Yang, and Z. Zhou, "Driving behavior evaluation model base on big data from internet of vehicles," *International Journal of Ambient Computing and Intelligence*, vol. 10, no. 4, pp. 78–95, 2019.

[27] Z. Deng, D. Chu, C. Wu, Y. He, and J. Cui, "Curve safe speed model considering driving style based on driver behaviour questionnaire," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 65, pp. 536–547, 2019.

[28] N. H. Gartner, C. J. Messer, and A. Rathi, "Traffic flow theory-a state-of-the-art report: revised monograph on traffic flow theory," Technical Report, US. Department of transportation, Washington, DC, USA, 2002.

[29] X. Qu, M. Yang, F. Yang, B. Ran, and L. Li, "An improved single-lane cellular automaton model considering driver's radical feature," *Journal of Advanced Transportation*, vol. 2018, Article ID 3791820, 10 pages, 2018.

[30] W. Knospe, L. Santen, A. Schadschneider, and M. Schreckenberg, "Towards a realistic microscopic description of highway traffic," *Journal of Physics A: Mathematical and General*, vol. 33, no. 48, 2000.

[31] H. Wang, "Experimental features and characteristics of speed dispersion in urban freeway traffic," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1999, no. 1, pp. 150–160, 2007.

WILEY | Hindawi

Research Article

Development of Fuzzy Level of Service Criteria for Bus Rapid Transit considering User Heterogeneities in China

Yueying Huo [ID],[1] Jianrong Liu,[2] Jian Zhang [ID],[3] and Xiaojuan Li[1]

[1]*Transportation Institute, Inner Mongolia University, 24 Zhaojun Rd, Hohhot, Inner Mongolia 010070, China*
[2]*School of Civil Engineering and Transportation, South China University of Technology, 381 Wushan Rd, Guangzhou, Guangdong Province 510640, China*
[3]*School of Transportation, Southeast University, 2 Southeast University Rd, Nanjing, Jiangsu Province 211189, China*

Correspondence should be addressed to Yueying Huo; hyy@imu.edu.cn

Level of service (LOS) analysis based on LOS criteria is essential for the planning, design, and operational evaluation of public transit. However, there are no systematic transit LOS criteria at present in China. Bus rapid transit (BRT) is receiving increasing attention worldwide. Therefore, this study addresses LOS criteria for BRT in China. Transit passengers are heterogeneous in their perceptions, needs, and behavior. The traditional hard LOS criteria have an inherent weakness, because of which the accuracy of an LOS analysis is limited. Thus, in this study, we initially conducted transit market segmentation to reduce heterogeneity and subsequently developed BRT fuzzy LOS criteria for different passenger groups. Using a smartphone-based transit travel survey system, we organized BRT passenger travel surveys on three BRT systems in China to collect data. Transit market segmentation was performed based on user perceptions; passengers were segmented into a calm passenger group and an anxious passenger group using the latent class model. Passenger arrival time, passenger wait time, and running speed of the bus were selected as service metrics to reflect the BRT's LOS. BRT fuzzy LOS criteria for the three service metrics in the case of both the calm and anxious passenger groups were developed using fuzzy C-means clustering. The LOS criteria for the two groups of passengers fit their psychological characteristics and reflected their personalized travel needs. Fuzzy LOS criteria can describe to what extent service metric values belong to the adjacent LOS categories via the use of membership. Thus, fuzzy LOS criteria can overcome the weakness of hard LOS criteria.

1. Introduction

Level of service (LOS) criteria using service metrics that classify them into several categories according to different thresholds can provide quantitative LOS analysis standards for transportation systems [1, 2]. LOS analysis based on LOS criteria is very essential for the planning, design, and operational evaluation of transportation facilities or services. Additionally, it is critical for the allocation of limited financial resources to competing transportation projects [1–3]. Transit capacity and quality of service manuals (TCQSMs) provide LOS criteria for public transit services. However, the LOS criteria in the TCQSMs are more suitable for North America, because that is where most of the research is conducted [4, 5]. In China, there are no systematic

transit LOS criteria at present. Therefore, it is necessary to develop transit LOS criteria specific to China in order to implement the LOS analysis.

The bus rapid transit (BRT) combines the efficiency and reliability of a rail service with the operating flexibility and low cost of a conventional bus service [6]. Currently, 168 cities worldwide have constructed BRT corridors and their total length has reached 4998 kilometers [7]. The topic of BRT LOS is receiving increasing attention, and we are carrying out research on it funded by the National Nature Science Foundation of China. Thus, we address the issue of BRT LOS criteria in China.

Transit passengers are heterogeneous [8]. Passenger behavior, need, and perception vary across different groups of passengers. Transit market segmentation, which defines

specific subgroups of passengers sharing similar demographic, psychographic, or behavioral characteristics, can reduce passenger heterogeneity [9, 10]. It permits transit providers to comprehend the special travel needs of certain subgroups so that they can cater to personalized travel requirements [9]. Thus, in this study, we will initially carry out transit market segmentation and subsequently develop BRT LOS criteria for different passenger groups.

LOS criteria in the TCQSMs or highway capacity manuals (HCMs) use clear-cut boundaries to divide multiple LOS categories [1, 2, 4, 5], in which a certain value of a service metric exclusively belongs to one LOS category, known as the hard LOS criteria in this study. An inherent weakness of the hard LOS criteria is that certain small changes in service metrics can cause a change in the LOS, while certain significant changes in service metrics may not lead to a variation in LOS. This limits the accuracy with which transportation facilities or services may be characterized [11, 12]. In order to overcome this weakness of the hard LOS criteria, we will try to develop fuzzy LOS criteria for the BRT in China.

The intent of this research work is to develop BRT fuzzy LOS criteria for passengers grouped using transit market segmentation in China. The research results can be used to implement transit LOS analysis and further guide the planning, design, and operational evaluation of transit services in China. The research methodology can act as a reference source for other regions in developing their own transit LOS criteria.

This paper is organized as follows. A literature review is first presented. Next, data collection, including the development of a transit travel survey system and BRT passenger travel surveys, is described. After this step, the methodology involved in a latent class model and fuzzy C-means clustering are introduced. Finally, the results from transit market segmentation and BRT fuzzy LOS criteria are analyzed and conclusions are summarized.

2. Literature Review

The LOS criteria provided by TCQSMs for bus transit with service metrics of frequency, service span, access, passenger load, on-time performance, headway adherence, and transit-auto travel time ratio are divided into multiple levels [4, 5]. Transit passengers are demographically, psychologically, and behaviorally heterogeneous [8]. However, the TCQSMs did not consider the heterogeneity of passengers and provided a set of general LOS criteria for all classes of passengers. Transit market segmentation can reduce passenger heterogeneity and permit transit providers to comprehend the special travel needs of certain subgroups so that they can cater to personalized travel requirements [9].

The most basic form of transit market segmentation is through social demographic and travel characteristics. Dell'Olio et al. stratified passengers by socioeconomic characteristics such as gender, age, and income. Further, models of service quality for different types of passengers were established using ordered probit models. This study demonstrated that the user perception of service quality

changed depending on the types of passengers [13]. Oña et al. addressed heterogeneity in user perceptions of service quality for railway service. Users were divided into different groups according to travel purpose, day of journey, and frequency of use. Based on questionnaire surveys conducted in Milan, northern Italy, models for different groups of users were established. This study again found that perceptions about service quality were different among various groups of users [14]. Using the data collected from questionnaire surveys in Kaohsiung, Cheng and Liu examined users' perceptions of intermodal inconvenience for the bicycle-transit service by utilizing the Rasch model. Various subgroups of users were defined based on socioeconomic and travel characteristics. This study demonstrated that users' perceived inconveniences differed according to gender, riding frequency, trip purpose, and environmental awareness [15]. Cheng and Chen evaluated public transportation service chains along the three dimensions of accessibility, mobility, and connectivity from the perspective of urban travelers. The analysis based on the Rasch model indicated that young passengers, environmentally conscious users, passengers using public transportation as their main transportation mode, and those who indulged in frequent sporting activities were better able to overcome difficulties during their trips [16]. Bordagaray et al. used ordered probit models to model bus transit quality in the city of Santander considering user and service heterogeneity. The perceived quality models were developed by introducing systematic variations in the users' perceptions to account for various factors such as age, income, and frequency of service use. The results show that heterogeneity is clearly present in the perception of service quality [17]. Passenger segmentation by social demographic and travel characteristics defines classes of passengers but not market segments. It gives rise to unbalanced segments where passengers do not share similar interests, needs, or locations, i.e., heterogeneity among users is still evident [9].

With the proliferation of automated data collection systems such as smart card automated fare collection (AFC), transit market segmentation can be performed with AFC data. Lathia et al. identified behavioral differences between passengers using AFC data. They applied agglomerative hierarchical clustering to segment passengers using temporal travel patterns. A number of algorithms based on historical travel time data were proposed for personalized trip time estimations for different types of passengers. By applying these algorithms to the data from the oyster fare collection system, London, it was empirically demonstrated that these algorithms outperformed both a nonpersonalized baseline computed from the data, as well as published travel times as currently offered by the transport authority [18]. There is a need for an integrated market segmentation method that incorporates both the spatial and behavioral features of individual transit passengers. Kieu et al. proposed a new algorithm named spatial affinity propagation (SAP) based on the classical affinity propagation algorithm (AP) to enable transit market segmentation with spatial-behavioral features. SAP was applied to a 40-day AFC dataset from New South Wales, Australia. This case study on a limited sample

size demonstrated that AP and SAP were better candidates for spatial passenger segmentation than K-means and complete-linkage hierarchical agglomerative clustering. Further, SAP was 52% more efficient in computation time than AP. The case study carried out on the entire dataset showed that SAP could deal with large-scale AFC data for transit market segmentation [9].

Some studies use the latent class model to conduct transit market segmentation using social demographic, travel, and behavioral characteristics as manifest variables. Oña et al. employed the latent class model to study how bus passengers could be stratified. Based on the data from customer satisfaction surveys in Granada, four groups of passengers were identified by the manifest variables of age, gender, travel reason, use frequency, possession of private vehicle, and ticket. Next, a Pearson correlation was applied for each group to derive the importance of the service quality attributes. It was discovered that the most important attributes for each group were different [10]. Qiao et al. addressed the issue of passenger market segmentation for high-speed railway service in China. The latent class model was used for passenger segmentation using age, gender, travel date, travel distance, mean time to get a ticket, and prepurchase time as manifest variables. Based on the ticket data of the Beijing-Shanghai high-speed railway, three distinct groups were identified via segmentation. They were the leisure passengers, planned passengers, and temporary passengers [19], respectively. Bellizzi et al. used the latent class model to examine the heterogeneity in the desired bus service quality from the current users' and potential users' perspective. Based on the preference data collected from stated preference surveys, current users were divided into three latent classes. These three classes accorded more importance to journey time (30% of users), comfort (60% of users), and fare (10% of users), respectively. However, potential users were segregated into two latent classes, and they conferred more importance to journey time (47% of users) and fare (53% of users) [20], respectively.

The TCQSMs provided hard LOS criteria for bus transit by the use of clear-cut boundaries in the division of various LOS categories [4, 5]. The HCMs provided hard LOS criteria for various transportation facilities from the intersection to the freeway [1, 2]. An inherent weakness of the hard LOS criteria is that some small changes in service metrics can cause a change in the LOS. On the other hand, even certain significant changes in service metrics may not lead to a change in LOS, which limits the accuracy with which transportation facilities or services may be characterized [11, 12]. By contrast, Fang et al. and Fang and Pecheux addressed fuzzy LOS criteria for signalized intersections. They developed a methodology using fuzzy C-means clustering to define fuzzy LOS criteria based on user perception. Using a user perception database of 100 subjects assessing 24 approaches in terms of estimated delay and rating of LOS, the fuzzy LOS criteria for signalized intersections were proposed. Compared with the hard LOS criteria of signalized intersections in the HCMs, fuzzy LOS criteria are capable of expressing the delays near the boundaries of each LOS category [11, 12].

In summary, because of the heterogeneity in transit passengers, we should perform transit market segmentation when studying transit LOS. The inherent weakness of the hard LOS criteria compelled us to develop fuzzy LOS criteria for transportation facilities or services. Therefore, we studied BRT fuzzy LOS criteria for the problem of passenger market segmentation in China.

3. User Perceptions

Transit passengers are heterogeneous and passengers' perceptions vary across different groups of passengers. The definition of LOS indicates that it needs to reflect user perceptions of quality of service provided by a facility or service. Thus, we conducted transit market segmentation based on user perceptions.

The journey of an individual taking a BRT bus includes traveling from the origin to the boarding station, waiting for a bus at the station, riding on the bus, and traveling from the alighting station to the destination. We selected a representative factor that influences user perceptions of quality of service in each subjourney. By doing so, user perceptions were defined as five-dimensional vectors of perceived arrival time, wait time, bus speed, passenger load, and departure time in this study. Perceived arrival time is the time taken by a passenger traveling from the origin to the boarding station as perceived by the passenger, which is different from the actual arrival time. Perceived wait time is the time a passenger spends waiting for a bus at a bus station according to his or her own perception, which differs from the actual wait time. Perceived bus speed is a passenger experience rating on the bus running speed while commuting in the bus. This is a categorical variable with a value of one expressing a rating of "very poor" and a value of five representing a rating of "very good." Perceived passenger load is a passenger experience rating on crowding encountered while riding on the bus, which is a categorical variable with a value of one expressing a rating of "very poor," and a value of five representing a rating of "very good." Note that in BRT passenger travel surveys, perceived bus speed and passenger load had a scale of one (very poor) to ten (very good). The ten-point scale was changed into a five-point scale in the analysis stage with the values one and two on the original scale integrated into the value one on the new scale. Similarly, the original values three and four were integrated into the value two and so on. The perceived departure time is the time taken by a passenger traveling from the alighting station to the destination according to his or her own perception, which is different from the actual departure time.

4. Data Collection

4.1. Transit Travel Survey System. We developed a smartphone-based transit travel survey system. This system consists of an app, a server, and a web interface. The app (Figure 1) interacts with users and collects and uploads the data to the server. The server stores the data and prepares them for the web interface. The web interface is the output

end of the system, from which the collected data are downloaded.

The app covers the entire trip of a passenger using the transit system and can record the complete travel information. Personal demographics, time elapsed, longitudes, and latitudes, that users present at each travel node, and user perceptions, can all be recorded. All the recorded data above are transmitted in real-time to the server. The server computes travel times in each subjourney and the bus speed, using real-time algorithms. Furthermore, it obtains the ride distance by invoking the Baidu map program. The server stores all of the data and transmits them to the web interface.

The collected data by the transit travel survey system include personal demographics, time spent by users at each travel node, times that the users spend in each subjourney, user perceptions, and other travel information. In particular, personal demographics contain gender, age, education, occupation, income, car ownership status, travel mode, BRT use frequency, and user's city of habitation. The time elapsed from the moment that a user sets out from the origin, arrives at the boarding station, boards a bus, alights from the bus, and arrives at the destination is collected. Arrival time, wait time, in-vehicle time, and departure time are all gathered. User perceptions, i.e., perceived arrival time, wait time, bus speed, passenger load, and departure time, are gathered. Other travel information such as travel purpose, bus route number, ride distance, and bus speed is collected as well.

4.2. BRT Passenger Travel Surveys. The Guangzhou BRT was opened in 2010 and has an average daily ridership of more than 850,000 passengers, higher than most metro and all light rail lines worldwide. It won the sustainable transport award in 2011 and is the only gold standard BRT in Asia. The Changzhou BRT was opened in 2008; its daily ridership exceeds 310,000 passengers, which accounts for 30% of the total public transit ridership. It won the Tien–Yow Jeme Civil Engineering Prize in 2010, making it the first public transit project to have won this prize. The Yichang BRT opened in 2015 and serves over 240,000 riders a day. Twenty percent of BRT riders previously drove a car or took a taxi. The city of Yichang won the sustainable transport award in 2016. We selected the Guangzhou, Changzhou, and Yichang BRTs to conduct passenger travel surveys using the transit travel survey system. The basic survey process includes recruiting participants and training them to use our cellphone app to collect the relevant data during daily travel.

Recruiting participants was the most difficult but crucial step. We tried several strategies to gather participants, such as by recruitment via an investigative firm, enrollment from WeChat groups, and enlistment via recommendations from acquaintances. However, these strategies could only bring together a small number of samples. Finally, we tried out a scheme wherein participants were recruited face-to-face at stations and on buses. Our experience with this technique proves that face-to-face communication with passengers at stations or on buses is the most effective way of recruiting participants. Four of us, acting as surveyors, entered the stations and randomly looked for passengers to converse with one-on-one. To begin with, we introduced ourselves, described our purpose in talking to them, and explained the task we were engaged in. Next, we invited passengers to participate in the survey. If the passenger was willing to participate, we added him/her to WeChat. This is very important because we had to teach them how to use the app through WeChat. We send the app's two-dimensional code and the reward both through WeChat as well. Meanwhile, we encouraged him/her to immediately download and install the app (1–2 minutes to complete such an install) by scanning the app's two-dimensional code at the station. If he/she has not installed the app when the bus arrives, we sent the app's two-dimensional code by WeChat and guided the person through the installation later. The participants recruited by a single surveyor in one day comprise a WeChat group. We reminded participants of each such group to use the app in daily travel and solve common problems that are encountered. The disadvantage of the aforementioned strategy is that usually the WeChat id of the passengers is not added by the time the bus arrives. This resulted in much wasted effort. Therefore, we attempted to recruit participants on the bus itself. We selected a bus that traveled on a route within the BRT system and randomly invited passengers to participate in the survey. We had ample time to communicate with passengers on the bus. However, surveyors became exhausted and succumbed to motion sickness owing to the rough terrain that the bus traveled over. Surveyors often had no energy to move back and forth to look for passengers after getting two or three participants and therefore, rarely sought out participants on the bus.

Those stations having large passenger volumes were selected to recruit participants. In the case of the Guangzhou BRT, we surveyed the stations of Gangding, Normal University and Jinan University, Shipai Qiao, and Tangxiacun. We also took route B1. The four surveyors representing each of the stations mentioned above worked from November 4 to 6, 2017, and about 65 participants were added. In the case of the Changzhou BRT, the stations of Wanfu Qiao, Huaide Road, Yanling Road, Lanling Road, and Renmin Road and Qinggong College were surveyed, and 100 participants were recruited over the period ranging from November 26 to 29, 2017. In the case of the Yichang BRT, the working stations were Wuyi Square, Pagoda River, Liujiadayan, and Lvluo Road. and 100 participants were enlisted from January 16 to 19, 2018. In these three cities, 265 participants were recruited in total.

After finding participants who were interested in the survey, data collection began. Participants took part in the survey for data collection using our app during daily BRT travel. A single data record was received for every participant who successfully used the app on a BRT trip. The data collection continued for one month in each city. Over the duration of the survey, participants were free to decide whether to continue their participation. Surveyors frequently encouraged them to use the app when taking the BRT.

From the participants' point of view, participation in survey took place as follows: (1) Surveyors invite passengers to participate in the BRT passenger travel survey. (2)

Figure 1: App interface (partial).

Passengers accept their invitation and become participants. (3) Participants download and install the app and register. (4) Prior to the formal survey, the participants take a pilot survey to be familiar with app. (5) During daily BRT travel, participants use the app over the entire trip to participate in the survey. They invoke the app and log in before they leave

the originating point of their journey. They operate the app following the prompts at each of the travel nodes that comprise the entire trip. The app runs in the background when no operations are being performed. (6) After completing the BRT travel route (i.e., there is no more use for the app), participants take a screenshot of the final app interface and send the screenshot to surveyors. (7) Surveyors judge whether the participants have used the app truly and correctly (i.e., whether the collected data are valid) by studying this final interface. (8) After completing a valid survey (i.e., correctly using the app on a single BRT trip), participants receive a reward of $1.50. Participants are free to end or continue their participation.

4.3. Sample Characteristics. The BRT passenger travel surveys produced a dataset with a sample size of 1304. Fifty-five percent of the sample was male. Ninety-five percent of the participants in the sample fell in the 19–30 year old age group or in the 31–50 year old age group. Eighty-three percent of the sample held a degree from a technical college or a bachelor's degree. Different occupations were equally represented in the sample. Sixty-five percent of the participants in the sample had a monthly income lower than $780. Sixty-three percent of the sample did not own a car. Most of the sample participants were frequent transit users. Among the participants in the sample, ninety-eight percent of them used the BRT for transit. Sixty-eight percent of the participants in the sample used the BRT every day. Sixty-seven percent of the sample stated that the purpose of travel was to commute, and twenty-five percent claimed recreation as the reason for travel. In general, the sample was, perhaps not surprisingly, composed mostly of young and middle-aged people, riders with a midlevel to upper level of education, low-income riders, riders without a car, and frequent transit users.

5. Methodology

We conducted transit market segmentation based on user perceptions. Because a user's perceived bus speed and passenger load are multivariate categorical data, we employed the latent class model to conduct BRT passenger segmentation.

Over the course of the entire journey of a passenger taking a BRT bus, the following factors influence user perceptions—passenger arrival time, wait time, bus speed, passenger departure time, facility condition (width, height, isolation, and cleanliness), availability of real-time bus arrival information, availability of shelter and benches, passenger load, driver attitude, and seating comfort. Of all the factors mentioned here, passengers most easily perceive those that are time related. Thus, passenger arrival time, passenger wait time, bus running speed, and passenger departure time were selected as service metrics to reflect BRT LOS. We developed fuzzy LOS criteria for BRT using fuzzy C-means clustering to deal with the ambiguity in the criteria.

5.1. Latent Class Model (LCM). The LCM seeks to stratify the cross-classification table of manifest variables by a latent variable that eliminates all confounding factors between the manifest variables [21]. The LCM approximates the observed joint distribution of the manifest variables as the weighted sum of a finite number R, denoting the number of latent classes, of constituent cross-classification tables [21]. The probability that an individual i in class r $(r = 1, 2, \ldots, R)$ produces a particular set of outcomes on the manifest variables, assuming local independence, is the product in the following equation:

$$f(Y_i; \pi_r) = \prod_{j=1}^{J} \prod_{k=1}^{K_j} \left(\pi_{jrk} \right)^{Y_{ijk}}, \tag{1}$$

where J denotes the number of manifest variables, $j = 1, 2, \ldots, J$; K_j denotes the number of categories (outcomes) of the jth manifest variable, $k = 1, 2, \ldots, K_j$; Y_{ijk} denotes the observed values of the manifest variables such that $Y_{ijk} = 1$ if an individual i $(i = 1, 2, \ldots, N)$ provides the kth outcome to the jth manifest variable, and $Y_{ijk} = 0$ otherwise; π_{jrk} denotes the conditional probability that an individual i in class r produces the kth outcome on the jth manifest variable. Within each class, for each manifest variable, $\sum_{k=1}^{K_j} \pi_{jrk} = 1$.

The probability density function across all classes is the weighted sum in the following equation:

$$P(Y_i \mid \pi, p) = \sum_{r=1}^{R} p_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} \left(\pi_{jrk} \right)^{Y_{ijk}}, \tag{2}$$

where p_r denotes the R mixing proportions that provide the weights in the weighted sum of the component tables, with $\sum_r p_r = 1$.

The parameters estimated by the LCM are p_r and π_{jrk}. The LCM is estimated by maximizing the log-likelihood function in the following equation:

$$\log L = \sum_{i=1}^{N} \ln \sum_{r=1}^{R} p_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} \left(\pi_{jrk} \right)^{Y_{ijk}}, \tag{3}$$

with respect to p_r and π_{jrk}; the expectation maximization (EM) algorithm is used for parameter estimation.

A priori, the number of the latent class R is unknown. Therefore, the goal of model selection is to find the model that is parsimonious and fits better to the study data. Model selection is conducted by fitting a complete independence model with $R = 1$ and iteratively increasing the number of latent class by one until a suitable fit has been achieved. The two most widely used criteria of model selection are the Bayesian information criterion (BIC) and the Akaike information criterion (AIC). Preferred models are those that minimize the values of the BIC and/or AIC [10].

5.2. Fuzzy C-Means Clustering (FCM). The FCM is a clustering algorithm that groups the data points in multidimensional space into a specific number of clusters.

Furthermore, every data point in the dataset belongs to every cluster according to a certain membership [22]. The purpose of FCM is to determine the cluster center $V = (v_1, v_2, \ldots, v_c)$ and the associated membership matrix $U = \{u_{ij}\}$. It is based on the minimization of the objective function in

$$J_m = \sum_{j=1}^{n} \sum_{i=1}^{c} \mu_{ij}^{m} \left\| x_j - v_i \right\|^2, \tag{4}$$

where x_j is the jth measured data point; n is the number of data points; v_i is the ith cluster center; c is the number of clusters; u_{ij} is the membership of x_j in cluster i; and m is a fuzziness index.

Fuzzy partitioning is carried out by utilizing an iterative optimization of the objective function with an adjustment of membership u_{ij} and cluster center v_i by using the following equations [22]:

$$\mu_{ij} = \frac{1}{\sum_{q=1}^{c} \left(\left\| x_j - v_i \right\| / \left\| x_j - v_q \right\| \right)^{2/(m-1)}}, \quad 1 \le i \le c, 1 \le j \le n, \tag{5}$$

$$v_i = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} x_j}{\sum_{j=1}^{n} \mu_{ij}^{m}}, \quad 1 \le i \le c. \tag{6}$$

This iteration will stop when $\max_{ij} (|u_{ij}^{(k+1)} - u_{ij}^{k}|) < \varepsilon$, where ε is a termination criterion.

6. Results

6.1. Results of Transit Market Segmentation. Based on the user perceptions data collected from BRT passenger travel surveys, we carried out transit market segmentation using LCM. Transit market segmentation aims to discover differentiated passenger groups according to perception or psychological factors. An assessment of whether the stratified passenger groups are significantly different was carried out based on the conditional probabilities (π_{jrk}) in LCM. In this research work, user perceptions include the perceived arrival time, wait time, bus speed, passenger load, and departure time. Perceived arrival, wait, and departure time are continuous variables and need to be converted into categorical variables.

Initially, all of these types of user perceptions were assumed to be manifest variables of the LCM to stratify passengers. However, the passenger groups stratified by assigning latent class number from one to ten exhibited no significant differences from one another. We also tried to combine some types of user perceptions as manifest variables and still did not find significantly different passenger groups. Next, all or some of these types of user perceptions were used as manifest variables, and passenger social-demographics such as gender, age, and education were used as covariates; latent class regression models were adopted to classify passengers; however, again, no different passenger groups were identified.

We created two new variables known as the class arrival time ratio and the class wait time ratio. The arrival time ratio is defined as perceived arrival time divided by actual arrival time. The class arrival time ratio is defined as a categorical variable; it assumes a value of unity if the arrival time ratio is smaller than, or equal to one, and it becomes two if the arrival time ratio is greater than one. The wait time ratio is defined as the perceived wait time divided by the actual wait time. The class wait time ratio is defined as a categorical variable that assumes a value of unity if the wait time ratio is smaller than or equal to one and a value of two if the wait time ratio is bigger than one.

Using the aforementioned two new variables along with perceived bus speed, passenger load, and departure time as manifest variables and passenger sociodemographics as covariates, latent class regression models were developed. Various combinations of these variables to stratify passengers did not lead to the discovery of significantly different passenger groups.

By gradually changing the manifest variables and the number of latent classes, distinct passenger groups were finally discovered. Using the class arrival time ratio, class wait time ratio, and perceived bus speed as manifest variables, models with latent classes of one to four were developed and the significantly different passenger groups were identified. The values of AIC and BIC for models with one to four classes were 5563.84 and 5593.29, 5505.49 and 5569.29, 5512.34 and 5610.50, and 5523.34 and 5655.85. The values of AIC and BIC for the two-class model were minimal, and thus, this model was adjudged the most appropriate and parsimonious model. Therefore, we stratified BRT passengers into two groups based on the accuracy of the fit results of LCM. The results of the estimated conditional probabilities and class probabilities of the two-class model are shown in Table 1.

Group 1 is composed of 60% of the sample data. All members of group 1 (100%) are passengers whose wait time ratio ≤ 1 (perceived wait time $\le$ wait time), 57% of the passengers are those whose arrival time ratio ≤ 1 (perceived arrival time $\le$ arrival time), and 66% of the passengers are those whose perceived bus speed is good. Group 2 consists of 40% of the sample data. In group 2, 90% of the passengers are those whose wait time ratio > 1 (perceived wait time $>$ wait time), 62% of the passengers are those whose arrival time ratio > 1 (perceived arrival time $>$ arrival time), and 62% of the passengers are those whose perceived bus speed is good. Passengers who are in a hurry and anxious to catch the bus always perceive wait time or arrival time to be longer than their actual values. On the contrary, passengers who are not in a hurry and are calm and relaxed may perceive wait time or arrival time to be shorter than their actual values. Therefore, group 1 was called the calm passenger group and group 2, the anxious passenger group. The calm passenger group (group 1) perceives bus speed to be slightly better than the anxious passenger group (group 2).

6.2. Results for Fuzzy LOS Criteria. Based on the service metrics (passenger arrival time, passenger wait time, and bus running speed), data collected from BRT passenger

TABLE 1: Estimated conditional and class probabilities with two latent classes.

Manifest variables	Categories	Group 1	Group 2
Class arrival time ratio	1 (arrival time ratio ≤ 1)	0.57	0.38
	2 (arrival time ratio > 1)	0.43	0.62
Class wait time ratio	1 (wait time ratio ≤ 1)	1.00	0.10
	2 (wait time ratio > 1)	0.00	0.90
Perceived bus speed	1 (very poor)	0.12	0.11
	2 (poor)	0.00	0.09
	3 (neutral)	0.22	0.18
	4 (good)	0.38	0.32
	5 (very good)	0.28	0.30
Estimated class probabilities		0.60	0.40

travel surveys and BRT fuzzy LOS criteria for the calm and the anxious passenger groups were developed using FCM. Initially, these three service metrics were divided into six clusters for each passenger group using FCM, following which, cluster centers and membership matrices were obtained. Next, membership functions for each service metric for each passenger group were derived based on the membership matrices. Subsequently, fuzzy LOS criteria for these three service metrics for each passenger group were developed by analyzing the membership functions.

Figure 2 shows the membership functions for the calm passenger group. Figures 2(a), 2(c), and 2(e) are the original membership functions of passenger arrival time, passenger wait time, and bus running speed, respectively. Figures 2(b), 2(d), and 2(f) are their respective approximated membership functions. The six different curves in each figure represent six different clusters, i.e., six levels of service, A–F. Figure 3 depicts the membership functions for the anxious passenger group, which include the original and approximated membership functions of passenger arrival time, passenger wait time, and bus running speed. Membership functions for passenger departure time are almost the same as those for the passenger arrival time; therefore, they have not been shown herein.

Tables 2 and 3 show BRT fuzzy LOS criteria for the calm passenger group, in which Table 2 displays ranges and centers of service metrics for fuzzy LOS categories, A–F, Table 3 displays ranges of service metrics for a single primary LOS category (membership value >0.5) and a single secondary LOS category (membership value <0.5). Tables 4 and 5 show BRT fuzzy LOS criteria for the anxious passenger group, which includes the ranges and centers for fuzzy LOS categories (Table 4) and the ranges for one primary category and one secondary LOS category (Table 5). As observed from Tables 2–5, for each fuzzy LOS category A–F, the boundary values of the passenger wait time for the anxious passenger group are smaller than those for the calm passenger group. Furthermore, the boundary values of the bus running speed for the anxious passenger group are higher than those for the calm passenger group. This result matches with the psychological characteristics of the two passenger groups. For example, the upper boundary values of the passenger wait time for the anxious passenger group are 2.4, 4.4, 7.3, 12.6, and 19.3,

respectively, while the corresponding values for the calm passenger group are 2.9, 5.4, 8.5, 13.2, and 23.1. The lower boundary values of the bus running speed for the anxious passenger group are 27.4, 21.8, 18.1, 14.9, 10.8, and zero, respectively, while the corresponding values for the calm passenger group are 24.9, 20.9, 17.4, 14.2, 10.4, and zero. Anxious passengers are always in a hurry and have higher anticipations of BRT service quality, i.e., they expect shorter wait times and higher bus speeds. By contrast, calm passengers are more tolerant and are able to accept longer wait times and slower bus speeds. Therefore, exactly as shown in Tables 2–5, anxious passenger groups have smaller boundary values of passenger wait time and higher boundary values of bus running speed than those of the calm passenger groups.

LOS criteria in the TCQSMs or HCMs use clear-cut boundaries for separating each LOS category, wherein a certain value of service metric exclusively belongs to one LOS category, which is called the hard LOS criteria in this study. An inherent weakness of the hard LOS criteria is that some small changes in service metrics can appear significant, while some big changes in service metrics can appear trivial, which limits the accuracy with which transportation facilities or services may be characterized [11, 12]. Specifically, the LOS changes from one category to the other when service metric values near the boundaries generate a small variation, while the LOS remains unchanged when service metric values display a large variation from the lower to the upper boundaries. For example, Das and Pandit established hard LOS criteria for passenger wait time for bus transit in India, where the ranges for LOS' A, B, C, D, and E, respectively, are 0 min, 0.1–4.0 min, 4.1–20.0 min, 20.1–35.0 min, and >35.0 min. The values of 0 min, 4.0 min, 20.0 min, and 35.0 min are the clear-cut boundaries for separating LOS' A, B, C, D, and E. The LOS changes from B to C when passenger wait time varies from 4.0 min to 4.1 min, while the LOS remains at C when passenger wait time varies from 4.1 min to 20.0 min [23]. As observed from Tables 2–5, fuzzy LOS criteria replace clear boundaries with fuzzy boundaries, i.e., a certain degree of overlap between adjacent LOS categories exists. Furthermore, the service metric values in the overlapping sections belong to adjacent LOS categories with varying memberships. The values of the memberships are derived from membership functions (Figures 2 and 3).

(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 2: Membership functions for the calm passenger group. (a, c, and e) The original membership functions; (b, d, and f) the approximated membership functions.

Consider the fuzzy LOS criteria of passenger wait time for the anxious passenger group (the second row in Table 4) as an example. The sections of 1.2–2.4 min, 2.4–4.4 min, 4.4–7.3 min, 7.3–12.6 min, and 12.6–19.3 min are the overlapping sections of LOS' A and B, B and C, C and D, D and E, and E and F, respectively. The values of passenger wait time in these overlapping sections belong to adjacent LOS categories with different memberships as shown in Figure 3(d). The value of 4.0 min belongs to LOS' B and C with memberships of 0.1 and 0.9. The value of 4.1 min belongs to LOS' B and C with memberships of 0.05 and 0.95. That is, the values of 4.0 min and 4.1 min belong to LOS' B and C simultaneously with different memberships, rather than being exclusively classified into B or C as in the hard LOS criteria. Fuzzy LOS criteria can accurately describe the extent to which service metric values belong to the adjacent LOS categories by the use of membership. Furthermore, the continuous changes in membership can reflect the LOS changes resulting from variations in the service metric values, thus permitting fuzzy LOS criteria to effectively overcome the weakness inherent in the hard LOS criteria.
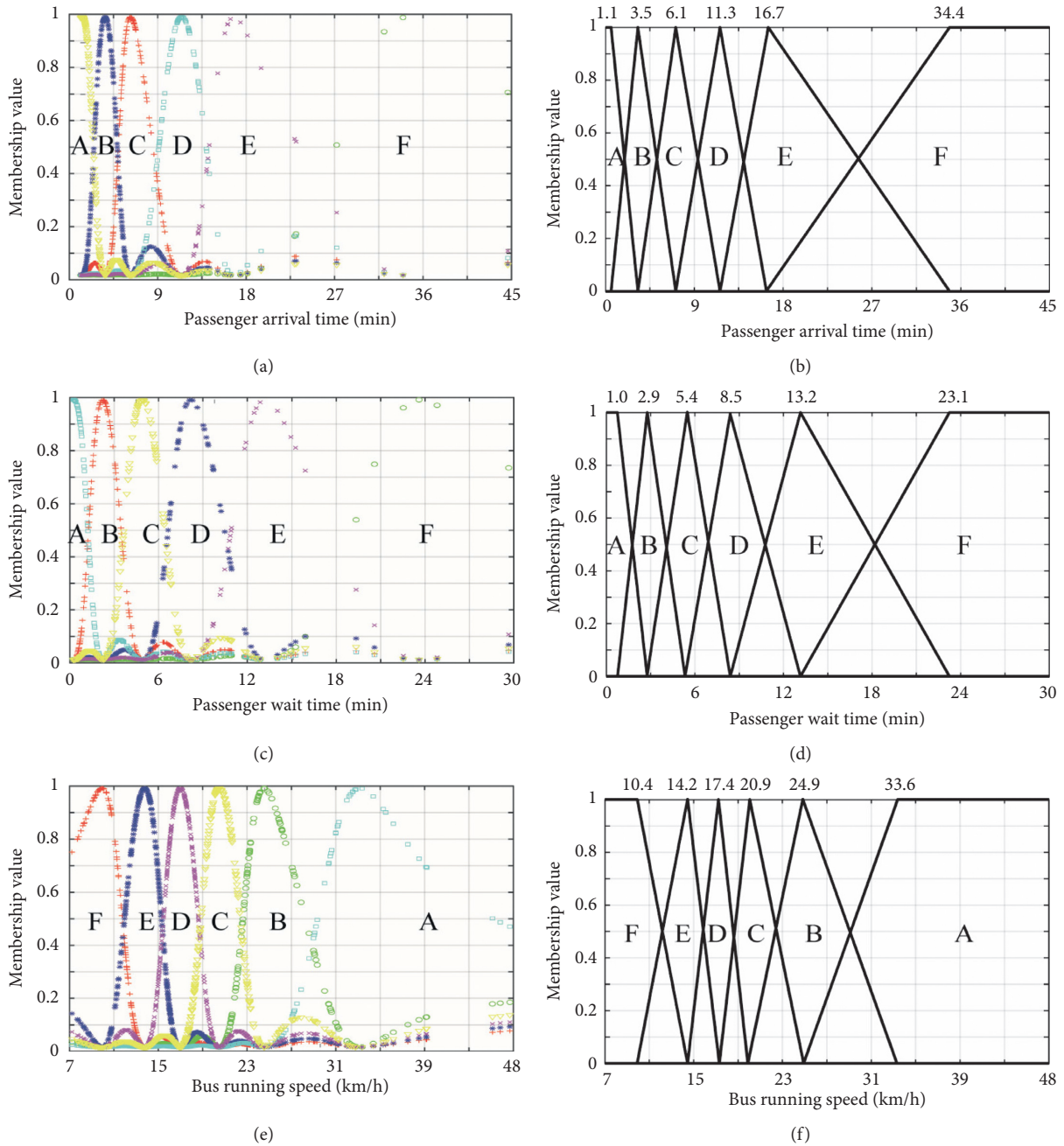
(a)

(b)

(c)

(d)

(e)

(f)
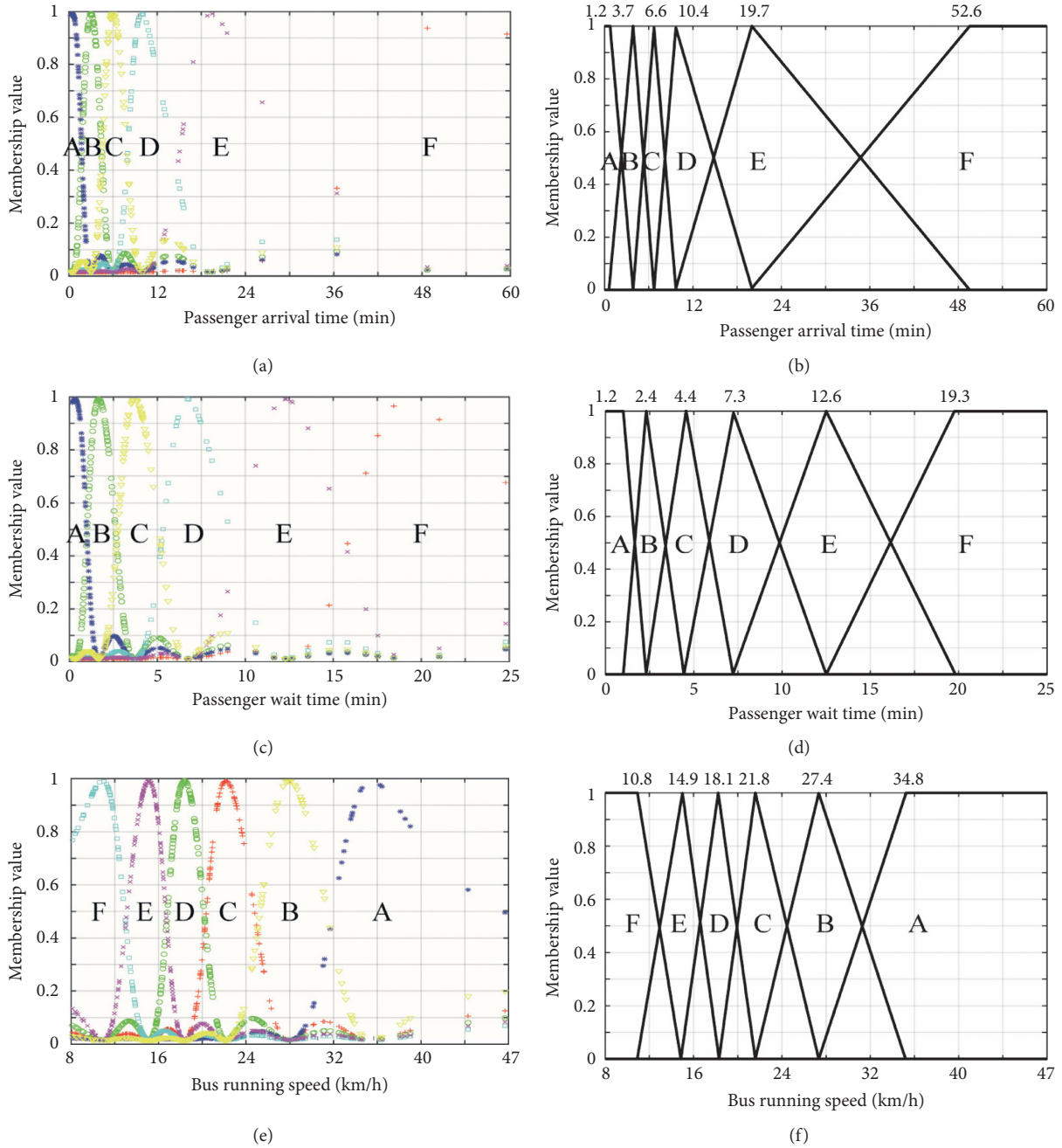
FIGURE 3: Membership functions for the anxious passenger group (a, c, and e) The original membership functions; (b, d, and f) the approximated membership functions).

TABLE 2: BRT fuzzy LOS criteria for the calm passenger group (ranges and centers of service metrics for fuzzy LOS categories).

Service metrics		Fuzzy LOS categories					
		A	B	C	D	E	F
Passenger arrival time (min)	Center	1.1	3.5	6.1	11.3	16.7	34.4
	Range	(0, 3.5)	(1.1, 6.1)	(3.5, 11.3)	(6.1, 16.7)	(11.3, 34.4)	(16.7, ∞)
Passenger wait time (min)	Center	1.0	2.9	5.4	8.5	13.2	23.1
	Range	(0, 2.9)	(1.0, 5.4)	(2.9, 8.5)	(5.4, 13.2)	(8.5, 23.1)	(13.2, ∞)
Bus running speed (km/h)	Center	33.6	24.9	20.9	17.4	14.2	10.4
	Range	(24.9, ∞)	(20.9, 33.6)	(17.4, 24.9)	(14.2, 20.9)	(10.4, 17.4)	(0, 14.2)

TABLE 3: BRT fuzzy LOS criteria for the calm passenger group (ranges of service metrics for one primary and one secondary LOS category).

Membership for fuzzy LOS categories		Service metrics		
		Passenger arrival time (min)	Passenger wait time (min)	Bus running speed (km/h)
$A > 0.5$	$B < 0.5$	(0, 2.3)	(0, 2.0)	(29.3, ∞)
$B > 0.5$	$A < 0.5$	(2.3, 3.5)	(2.0, 2.9)	(24.9, 29.3)
	$C < 0.5$	(3.5, 4.8)	(2.9, 4.2)	(22.9, 24.9)
$C > 0.5$	$B < 0.5$	(4.8, 6.1)	(4.2, 5.4)	(20.9, 22.9)
	$D < 0.5$	(6.1, 8.7)	(5.4, 7.0)	(19.2, 20.9)
$D > 0.5$	$C < 0.5$	(8.7, 11.3)	(7.0, 8.5)	(17.4, 19.2)
	$E < 0.5$	(11.3, 14.0)	(8.5, 10.9)	(15.8, 17.4)
$E > 0.5$	$D < 0.5$	(14.0, 16.7)	(10.9, 13.2)	(14.2, 15.8)
	$F < 0.5$	(16.7, 25.6)	(13.2, 18.2)	(12.3, 14.2)
$F > 0.5$	$E < 0.5$	(25.6, ∞)	(18.2, ∞)	(0, 12.3)

TABLE 4: BRT fuzzy LOS criteria for the anxious passenger group (ranges and centers of service metrics for fuzzy LOS categories).

Service metrics		Fuzzy LOS categories					
		A	B	C	D	E	F
Passenger arrival time (min)	Center	1.2	3.7	6.6	10.4	19.7	52.6
	Range	(0, 3.7)	(1.2, 6.6)	(3.7, 10.4)	(6.6, 19.7)	(10.4, 52.6)	(19.7, ∞)
Passenger wait time (min)	Center	1.2	2.4	4.4	7.3	12.6	19.3
	Range	(0, 2.4)	(1.2, 4.4)	(2.4, 7.3)	(4.4, 12.6)	(7.3, 19.3)	(12.6, ∞)
Bus running speed (km/h)	Center	34.8	27.4	21.8	18.1	14.9	10.8
	Range	(27.4, ∞)	(21.8, 34.8)	(18.1, 27.4)	(14.9, 21.8)	(10.8, 18.1)	(0, 14.9)

TABLE 5: BRT fuzzy LOS criteria for the anxious passenger group (ranges of service metrics for one primary and one secondary LOS category).

Membership for fuzzy LOS categories		Service metrics		
		Passenger arrival time (min)	Passenger wait time (min)	Bus running speed (km/h)
$A > 0.5$	$B < 0.5$	(0, 2.5)	(0, 1.8)	(31.1, ∞)
$B > 0.5$	$A < 0.5$	(2.3, 3.7)	(1.8, 2.4)	(27.4, 31.1)
	$C < 0.5$	(3.7, 5.2)	(2.4, 3.4)	(24.6, 27.4)
$C > 0.5$	$B < 0.5$	(5.2, 6.6)	(3.4, 4.4)	(21.8, 24.6)
	$D < 0.5$	(6.6, 8.5)	(4.4, 5.9)	(20.0, 21.8)
$D > 0.5$	$C < 0.5$	(8.5, 10.4)	(5.9, 7.3)	(18.1, 20.0)
	$E < 0.5$	(10.4, 15.1)	(7.3, 10.0)	(16.5, 18.1)
$E > 0.5$	$D < 0.5$	(15.1, 19.7)	(10.0, 12.6)	(14.9, 16.5)
	$F < 0.5$	(19.7, 36.2)	(12.6, 16.0)	(12.9, 14.9)
$F > 0.5$	$E < 0.5$	(36.2, ∞)	(16.0, ∞)	(0, 12.9)

7. Conclusions

This study presents a framework to develop fuzzy LOS criteria for bus transit considering passengers' heterogeneity. Further, it proposes BRT fuzzy LOS criteria for different passenger groups in China. A smartphone-based transit travel survey system was created, with which BRT passenger travel surveys were conducted in Guangzhou, Changzhou, and Yichang, and the data for user perceptions and service metrics were collected. Transit market segmentation was performed based on user perceptions, and passengers were segmented into the calm passenger group and the anxious passenger group using the LCM. Passenger arrival time, passenger wait time, and bus running speed were selected as service metrics to reflect BRT LOS. BRT fuzzy LOS criteria for the calm passenger group and the anxious passenger group, respectively, were proposed using fuzzy C-means clustering.

The TCQSMs did not consider passengers' heterogeneity and provided a set of general LOS criteria for bus transit. However, this research work examines heterogeneity by passenger market segmentation and proposes BRT LOS criteria for anxious and calm passenger groups. It was discovered that the anxious passenger group had lower boundary values of passenger wait time and higher boundary values of bus running speed than those of the calm passenger group, which fits the psychological characteristics of the two groups. Simultaneously, it demonstrates that the LOS criteria in this study can better reflect the personalized travel needs of certain passenger groups.

The TCQSMs provided hard LOS criteria for bus transit. An inherent weakness of the hard LOS criteria is that some small changes in service metrics can result in a change of LOS, while some significant changes in service metrics may not trigger such changes. However, this study proposes fuzzy LOS criteria for BRT in China. Fuzzy LOS criteria can

properly describe the extent to which service metric values belong to the adjacent LOS categories via the use of membership, so that they can overcome the inherent weakness of the hard LOS criteria.

It was quite difficult to recruit participants owing to the requirement that the transit travel survey system app has to be installed on the participants' handheld device. Therefore, the sample size used in this study was small. This is one limitation in the exploration of potential differences in perceptions between different cities. Our study focused on the heterogeneity in passenger perception. In the future, the heterogeneity both in passenger behavior and perception will be studied to conduct transit market segmentation.

Data Availability

The data used to support the findings of this study are available in supplementary information files.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this research work.

Acknowledgments

Supplementary Materials

The supplementary descriptions for the supplementary files. (1) **The descriptions for the file of mydata.csv**. The variables in the file of mydata.csv include gender, age, education, occupation, income, car, travel_mode, brt_freq, travel_purpose, arrival_time, arrival_time_p, wait_time, wait_time_p, speed_p, crowd_p, departure_time, departure_time_p, overall_p, young, middle_age, high_school, university, student, worker, low_income, middle_high_income, work, home, recreation, arrival_time_ratio, class_arrival_time_ratio, wait_time_ratio, and class_wait_time_ratio. The introductions for each variable are as follows. For gender, "0" represents female and "1" represents male. For age, "1" represents teenager and "2" represents middle-aged person. For education, "1" represents high school and below and "2" represents higher education. For occupation, "1" represents self-employed, "2" represents other, "3" represents enterprise staff, "4" represents students, "5" represents unemployed persons, and "6" represents civil servant. For income, "1" represents under 3000 yuan, "2" represents 3000–5000 yuan, "3" represents 5000–7000 yuan, and "4" represents over 7000 yuan. For car, it refers to if there is a private car, "0" represents "no," and "1" represents "yes." For travel_mode, "1" represents transit and "0" represents other. For brt_freq, it means use frequency of BRT, "1" represents occasional ride, "2" represents 1-2 days per week, "3" represents 3-4 days per week, "4" represents weekdays, and "5" represents every day. For travel_purpose, "1"

represents work travel, "2" represents going home, and "3" represents other. arrival_time means the time a passenger traveling from origin to boarding station. arrival_time_p, which is the perceived arrival time, is the time a passenger travels from the origin to the boarding station perceived from his own perspective. wait_time is the time a passenger spends in waiting at a bus at the station. wait_time_p is the perceived waiting time, which is the time a passenger spends in waiting for a bus at the station perceived by himself. speed_p is the bus speed perception, which is a passenger's experience rating on bus running speed generated during riding on the bus on a scale of 1 to 10 with 1 representing very poor and 10 representing very good. crowd_p is the passenger load perception, which is a passenger experience rating on crowding encountered during riding on the bus on a scale of 1 to 10 with 1 representing very poor and 10 representing very good. departure_time is the time a passenger traveling from alighting station to destination. departure_time_p is the perceived departure time, which is the time a passenger travels from a lighting station to the destination perceived by himself. overall_p means a passenger experience rating on operational condition of the whole bus travel from the origin to the destination on a scale of 1 to 10 with 1 representing very poor and 10 representing very good. Young means if a passenger is a teenager, "1" represents "yes" and "0" represents "no." middle_age means that if a passenger is a middle-aged person, "1" represents "yes" and "0" represents "no." high_school means if the education level of a passenger is high school, "1" represents "yes" and "0" represents "no." University means if a passenger received higher education, "1" represents "yes" and "0" represents "no." Student means if a passenger is a student, "1" represents "yes" and "0" represents "no." Worker means if a passenger is a worker, "1" represents "yes" and "0" represents "no." low_income means if a passenger belongs to low-income group, "1" represents "yes" and "0" represents "no." middle_high_income means if a passenger belongs to middle- or high-income group, "1" represents "yes" and "0" represents "no." Work means if travel purpose is work travel, "1" represents "yes" and "0" represents "no." Home means if travel purpose is going home, "1" represents "yes" and "0" represents "no." Recreation means if travel purpose is recreation, "1" represents "yes" and "0" represents "no." Arrival_time_ratio (i.e., arrival time ratio) is defined as the perceived arrival time divided by arrival time. Class_arrival_time_ratio (i.e., class arrival time ratio) is defined as a categorical variable with it being 1 if arrival time ratio is smaller than or equal to 1 and it being 2 if arrival time ratio is bigger than 1. Wait_time_ratio (i.e., wait time ratio) is defined as the perceived wait time divided by wait time. Class_wait_time_ratio (i.e., class wait time ratio) is defined as a categorical variable with it being 1 if wait time ratio is smaller than or equal to 1 and it being 2 if wait time ratio is bigger than 1. (2) **The descriptions for the file of mydata_class.csv**. Comparde with the file of mydata.csv, the file of mydata_class.csv adds a variable of "class." The variables in the file of mydata_class.csv include gender,

age, education, occupation, income, car, travel_mode, brt_freq, travel_purpose, arrival_time, arrival_time_p, wait_time, wait_time_p, speed_p, crowd_p, departure_time, departure_time_p, overall_p, young, middle_age, high_school, university, student, worker, low_income, middle_high_income, work, home, recreation, arrival_time_ratio, class_arrival_time_ratio, wait_time_ratio, class_wait_time_ratio, and class. For the same variables in these two files, they have the same meanings. BRT passengers are stratified into two classes using the latent class model in this paper. The variable of "class" represents the two stratified classes with "1" and "2." (*Supplementary Materials*)

References

[1] Transportation Research Board of the National Academies, *Highway Capacity Manual*, Transportation Research Board, Washington, DC, USA, 2000.

[2] Transportation Research Board of the National Academies, *Highway Capacity Manual*, Transportation Research Board, Washington, DC, USA, 2010.

[3] K. Choocharukul, K. C. Sinha, and F. L. Mannering, "User perceptions and engineering definitions of highway level of service: an exploratory statistical comparison," *Transportation Research Part A: Policy and Practice*, vol. 38, no. 9-10, pp. 677–689, 2004.

[4] Transportation Research Board of the National Academies, "Transit capacity and quality of service manual," Transportation Research Board, Washington, DC, USA, TCRP Report 100, 2nd edition, 2003.

[5] Transportation Research Board of the National Academies, "Transit capacity and quality of service manual," Transportation Research Board, Washington, DC, USA, TCRP Report 165, 3rd edition, 2013.

[6] R. Cervero and C. D. Kang, "Bus rapid transit impacts on land uses and land values in seoul, Korea," *Transport Policy*, vol. 18, no. 1, pp. 102–116, 2011.

[7] Bus rapid transit-centre of excellence ," 2019, http://www.brt.cl/.

[8] L. M. Kieu, A. Bhaskar, and E. Chung, "Passenger segmentation using smart card data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1537–1548, 2015.

[9] L. M. Kieu, Y. Ou, and C. Cai, "Large-scale transit market segmentation with spatial-behavioral features," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 97–113, 2018.

[10] R. D. Oña, G. López, F. J. D. D. L. Rios, and J. D. Oña, "Cluster Analysis for diminishing heterogeneous opinions of service quality public transport passengers," *Procedia-Social and Behavioral Sciences*, vol. 162, pp. 459–466, 2014.

[11] C. F. Fang, L. Elefteriadou, K. K. Pecheux, and M. T. Pietrucha, "Using fuzzy clustering of user perception to define levels of service at signalized intersections," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 657–663, 2003.

[12] C. F. Fang and K. K. Pecheux, "Fuzzy data mining approach for quantifying signalized intersection level of services based on user perceptions," *Journal of Transportation Engineering*, vol. 135, no. 6, pp. 349–358, 2009.

[13] L. Dell'Olio, A. Ibeas, and P. Cecín, "Modelling user perception of bus transit quality," *Transport Policy*, vol. 17, no. 6, pp. 388–397, 2010.

[14] J. D. Oña, R. D. Oña, L. Eboli, and G. Mazzulla, "Heterogeneity in perceptions of service quality among groups of railway passengers," *International Journal of Sustainable Transportation*, vol. 9, no. 8, pp. 612–626, 2015.

[15] Y.-H. Cheng and K.-C. Liu, "Evaluating bicycle-transit users' perceptions of intermodal inconvenience," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 10, pp. 1690–1706, 2012.

[16] Y.-H. Cheng and S.-Y. Chen, "Perceived accessibility, mobility, and connectivity of public transportation systems," *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 386–403, 2015.

[17] M. Bordagaray, L. Dell'Olio, A. Ibeas, and P. Cecín, "Modelling user perception of bus transit quality considering user and service heterogeneity," *Transportmetrica A: Transport Science*, vol. 10, no. 8, pp. 705–721, 2014.

[18] N. Lathia, C. Smith, J. Froehlich, and L. Capra, "Individuals among commuters: building personalised transport information services from fare collection systems," *Pervasive and Mobile Computing*, vol. 9, no. 5, pp. 643–664, 2013.

[19] K. Qiao, P. Zhao, and J.-X. Wen, "Passenger market segmentation of high-speed railway based on latent class model," *Journal of Transportation Systems Engineering and Information Technology*, vol. 17, no. 2, pp. 28–34, 2017.

[20] M. G. Bellizzi, L. Dell'Olio, L. Eboli, and G. Mazzulla, "Heterogeneity in desired bus service quality from users' and potential users' perspective," *Transportation Research Part A: Policy and Practice*, vol. 132, pp. 365–377, 2020.

[21] D. A. Linzer and J. B. Lewis, "poLCA: an R package for polytomous variable latent class analysis," *Journal of Statistical Software*, vol. 42, no. 10, pp. 1–29, 2011.

[22] S. Askari, N. Montazerin, and M. H. F. Zarandi, "Generalized possibilistic fuzzy C-means with novel cluster validity indices for clustering noisy data," *Applied Soft Computing*, vol. 53, pp. 262–283, 2017.

[23] S. Das and D. Pandit, "Determination of level-of-service scale values for quantitative bus transit service attributes based on user perception," *Transportmetrica A: Transport Science*, vol. 11, no. 1, pp. 1–21, 2015.

WILEY | Hindawi

Research Article

Sizing Up Transport Poverty Alleviation: A Structural Equation Modeling Empirical Analysis

Li Jiang, Huiying Wen [ID], and Weiwei Qi [ID]

School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, Guangdong, China

Correspondence should be addressed to Weiwei Qi; ctwwqi@scut.edu.cn

The aim of this paper was to systematically obtain the poverty reduction effects and diversified poverty alleviation paths of road infrastructure, to assist in the planning of transport poverty alleviation and rural revitalization in the concentrated contiguous poverty-stricken areas. The measurement methods for impact effects and paths of transport poverty alleviation have been scientifically proposed based on methods of transportation economics and transportation geography. Firstly, an overview of the interactive mechanism and processes by which road infrastructure investment influences poverty reduction has been offered, and the characteristics of poverty space and spatial coupling result of road infrastructure have been systematically obtained. The results show that about 70% of the district counties' road infrastructure and poverty rate are in a state of spatial coupling imbalance; the coordinated synchronous type is mainly distributed along the road network. Secondly, the structural equation model system has been formulated with variables that reflect transportation input in adjacent geographical units to consider spatial spillover effects. The results show that the direct poverty reduction effect of road infrastructure (0.105) is much lower than the indirect poverty reduction effect (0.830). Thirdly, empirical analysis at regional level of the concentrated contiguous poverty-stricken areas in China has been conducted. As a result, while addressing the limitations of previous studies, the poverty alleviation path that has been proposed also aims to catalyze actions to reduce the transport-related exclusion in poverty-stricken areas caused by the lack of access to basic facilities.

1. Background

Poverty is a global development problem and a practical problem. Antipoverty has become an important issue facing all mankind, especially developing countries, where poverty has seriously restricted the coordinated development of social economy [1]. One of the largest and most comprehensives surveys about poverty published to date is the "Voices of the Poor," which summarized 40,000 experiences of poor people from 50 different countries around the world. The findings of this report assert that poverty is perceived as consisting of many interlocking dimensions, in which lack of access to basic infrastructure, rural roads, and transportation is frequently pointed out as a remarkable factor [2]. In addition, evidence recently published by the World Bank has shown that the primary difference between those who have escaped chronic poverty and those are still trapped in it is

not income, but access to essential services [3]. Transport is one of the main sectors; the Asian Development Bank (ADB) supports a reflection of the important role that transport plays in enabling economic development and poverty reduction [4, 5].

In the realm of transportation research, road infrastructure is frequently cited as a critical catalyst for regional social transformation and economic growth; it can generate payoffs for local development [6, 7]. However, how the payoffs come into effect remains uncertain. Polzin categorized the impacts of road infrastructure on economy development into three types: direct, indirect, and secondary [8]. The direct impacts are those related to the improved service and accessibility; the indirect impacts are those related to mediating community responses and policy change; the secondary impacts are those related to social perception and individual behavioral changes. In this context, it is

worthwhile to examine the pathways for the influence of the road infrastructure on the local economy and poor reduction using quantitative measurements.

As a multivariate statistical technique, structural equation modeling (SEM) can discover the direct or indirect relationships between exploratory variables and a dependent variable and then visualize the causal influencing pathways and medicating effect [9, 10]. Previous pilot studies have indicated that SEM is suitable for social and economic studies, and it can deal with the problem of causality between public capital and output. However, rarely have studies applied SEM to investigate the relationships between the road infrastructure and economic development, as well as the transport poverty alleviation. The primary objective of this paper is to explore the relationships between the road infrastructure and regional economic development using SEM. In particular, analysis is conducted in the concentrated contiguous destitute areas of Yunnan in southwestern inland China.

The concentrated contiguous destitute areas of Yunnan located in southwestern China are shown in Figure 1. Among the 14 concentrated contiguous destitute areas in China, there are four in Yunnan, which are composed of 91 cities/counties. These regions, respectively, account for 65.89% and 12.50% of the total number of administrative units at the county level and the total number of contiguous areas in China.

The concentrated contiguous destitute areas of Yunnan are among the most underdeveloped hilly regions in China. Their socioeconomic conditions and road infrastructure construction are significantly lower than the national average. This region provides useful cases to study the pathway for the influence of the road infrastructure on the local economy among poverty-stricken areas with development.

Under these circumstances, this study aims to provide a comprehensive empirical evaluation of the impact that transportation development has on economic growth and poverty reduction, which can assist the planning of traffic poverty alleviation and rural revitalization in the contiguous poverty-stricken areas of Yunnan (in China). This paper differs from existing research in several aspects. First, an overview of the interactive mechanism and processes has been offered by which road infrastructure investment influences poverty reduction and the characteristics of poverty space and spatial coupling result of road infrastructure have been systematically obtained. Second, a structural equation model (SEM) has been built, which comprehensively considers the mutual relationship between transportation investment and poverty reduction; meanwhile, social capital, material capital, and geographical capital have been added as endogenous variables in the model system. Third, the model system has been formulated with variables that reflect transportation input in adjacent geographical units to consider spatial spillover effects. Fourth, empirical analysis at regional level of the concentrated poverty-stricken areas in China has been conducted. The poverty alleviation path that has been proposed aims to catalyze actions to reduce the transport-related exclusion in poverty-stricken areas caused by the lack of access to basic facilities (e.g., health, education, financial, and ecological facilities).

The remainder of this paper is structured as follows: Section 2 theorizes about the transport-region economic relationship. Then, the data sources and types are illustrated in Section 3, and the evaluation methods between road infrastructure and poverty rate are also applied. Section 4 mainly discusses the empirical results of research. Finally, this paper ends with some conclusions and discussion in Section 5.

2. Theorizing the Transport-Region Economic Relationship

2.1. Literature Review. Previous studies linking the local economy to road infrastructure can be traced back to the last century. The potential contributions of the transportation sector to poverty reduction have been investigated in several countries and institutions since the late 1960s [11–20].

In recent years, the conceptualization and measurement of transport disadvantages have garnered even more attention, particularly from academics and decision-makers of the developed countries, for being considered as a key driver of social inclusion and social justice [21–26]. This phenomenon has been illustrated by a number of studies pointing out that accessibility constraints tend to deepen sociospatial inequalities leading to multidimensional deprivations and, eventually, poverty traps [19, 27]. In this sense, accessibility, as the ultimate goal of most transportation, has been also proposed to be understood and planned as a fundamental human capability since it plays a central role in enabling people to meet their needs and promote healthy human flourishing [28–30]. Moreover, in spite of the increasing body of research dedicated to evaluating access to house facilities and healthcare services in rural China or low and middle income countries [31, 32], only few notable exceptions have been dedicated to quantitative evaluation of the overall accessibility poverty in such context [33–35].

Existing literature attempts to reveal the transport economic mechanisms using an econometric approach. These researches demonstrated the economic impact of transportation infrastructure investment through three pathways: (1) enhancing overall production productivity [36, 37], (2) accelerating technological spillovers across economic units [38, 39], and (3) reducing trade costs and enlarging sale spatial scope [40, 41].

Another branch of literature employs the classical statistical method to quantify the association between transportation infrastructure and the regional economy. An empirical analysis [42] showed that the impact of highway construction on population growth was weak in rural areas and strong in suburban areas. Guo proved that highway development, railway development, and railway investment all contributed positively to economic growth [43]. Iacono and Levinson pointed out that local road infrastructure and employment have no significant associations [7]. Yu found that road infrastructure improvements alone were not sufficient to promote economic growth at the national and

(a)   (b)
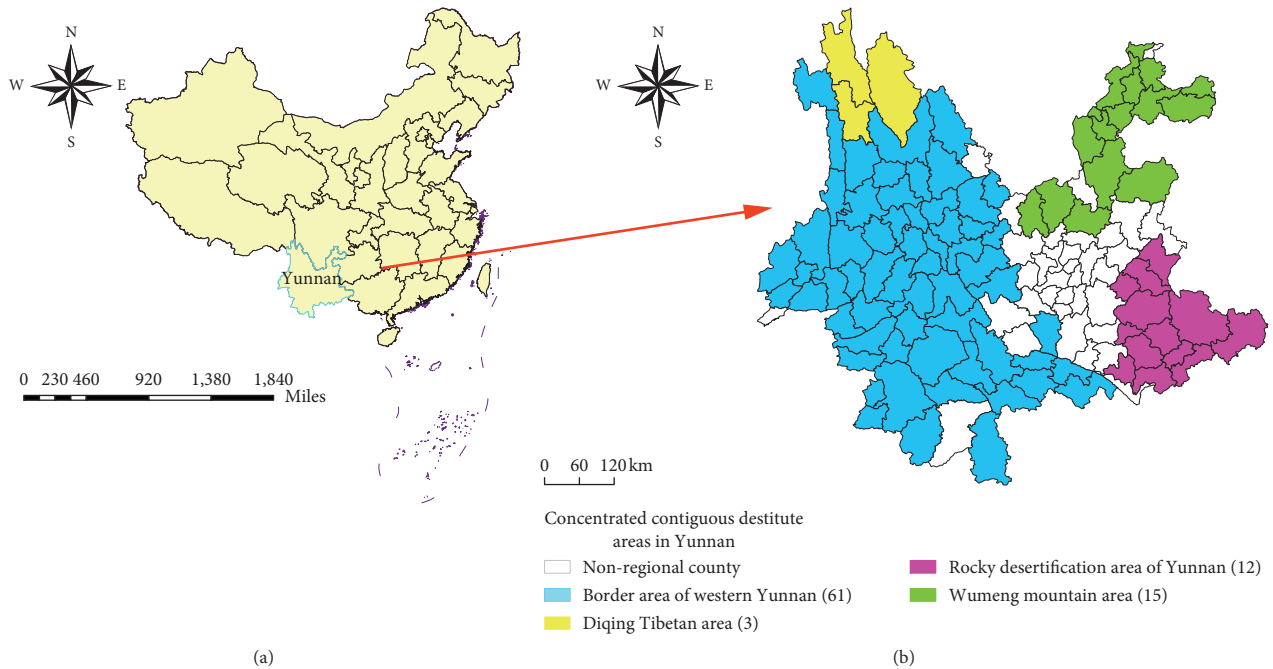
FIGURE 1: Location of the concentrated contiguous destitute areas of Yunnan, China.

regional levels in China [44]. Most of the above reviewed studies were case based and only show a certain specific pathway that explains the transport-region economic relationship.

These studies provide critical references for theorizing the transport-region economic relationship. Nevertheless, the studies on this subject are not comprehensive enough in geographical terms and have not covered most of the regions in developing countries where poverty is widespread [36, 45], especially in the contiguous poverty-stricken areas. The pathways for the influence of the road infrastructure on the local economy should be very complex. Poverty-stricken areas should be identified in order to compare the existing results of studies in developed regions. However, few efforts have focused on these areas and aspects. Additionally, no specific conceptual framework has been proposed to demonstrate the transport-region economic relationship with regard to road infrastructure construction.

2.2. Theoretical Framework. A growing body of empirical literature indicates that road infrastructure plays an important role in regional economic development, and it can generate payoffs for local development. However, there is uncertainty as to how the payoffs come into effect. Perfect road transportation service would positively increase the employment scale and income of poverty-stricken areas, influence the spatial location and layout of the market, and realize the effect of spatial agglomeration, while poor transportation infrastructure service may cause issues within the economic and social welfare system. The effects can be divided into two aspects: one is direct investment drive, and the other is indirect economic spillover. Figure 2 offers a view of the mechanism and processes by which road infrastructure influences economic development and alleviates poverty [46].

On one hand, improving road transportation investment and service can directly reduce the production cost and improve the productivity of relevant industries in poor areas. It is also known as "cost saving effect," which promotes total factors productivity growth (TFP) and the rate of economic growth in the poor areas. At the same time, road infrastructure investment can significantly reduce transport and travel costs, improve the accessibility and convenience of poor areas, and bring about multiplier effect of economic production activities.

On the other hand, the indirect effects of road infrastructure on economic growth and poverty alleviation can be categorized into three aspects: spatial spillover effect, spatial agglomeration and technical diffusion effect, and attraction effect. Road infrastructure has the characteristics of network and externality. Positive spatial spillovers take place when connectivity and accessibility in the surrounding areas are improved, those would then lead to the immigration of knowledge, technology, and resources to surrounding areas and strengthen attraction in the poor regions. In addition, the attraction effect would promote connectivity and improve resource allocation. The three effects and mechanisms influence the evolution and adjustment of regional spatial structure, thus promoting the development of regional economic. The "trickle-down effect" of economic growth can indirectly improve the quality of social services, such as medical care, education, and banking in poor areas and make it easier for the poor to enjoy facilities and services, so as to reduce poverty.
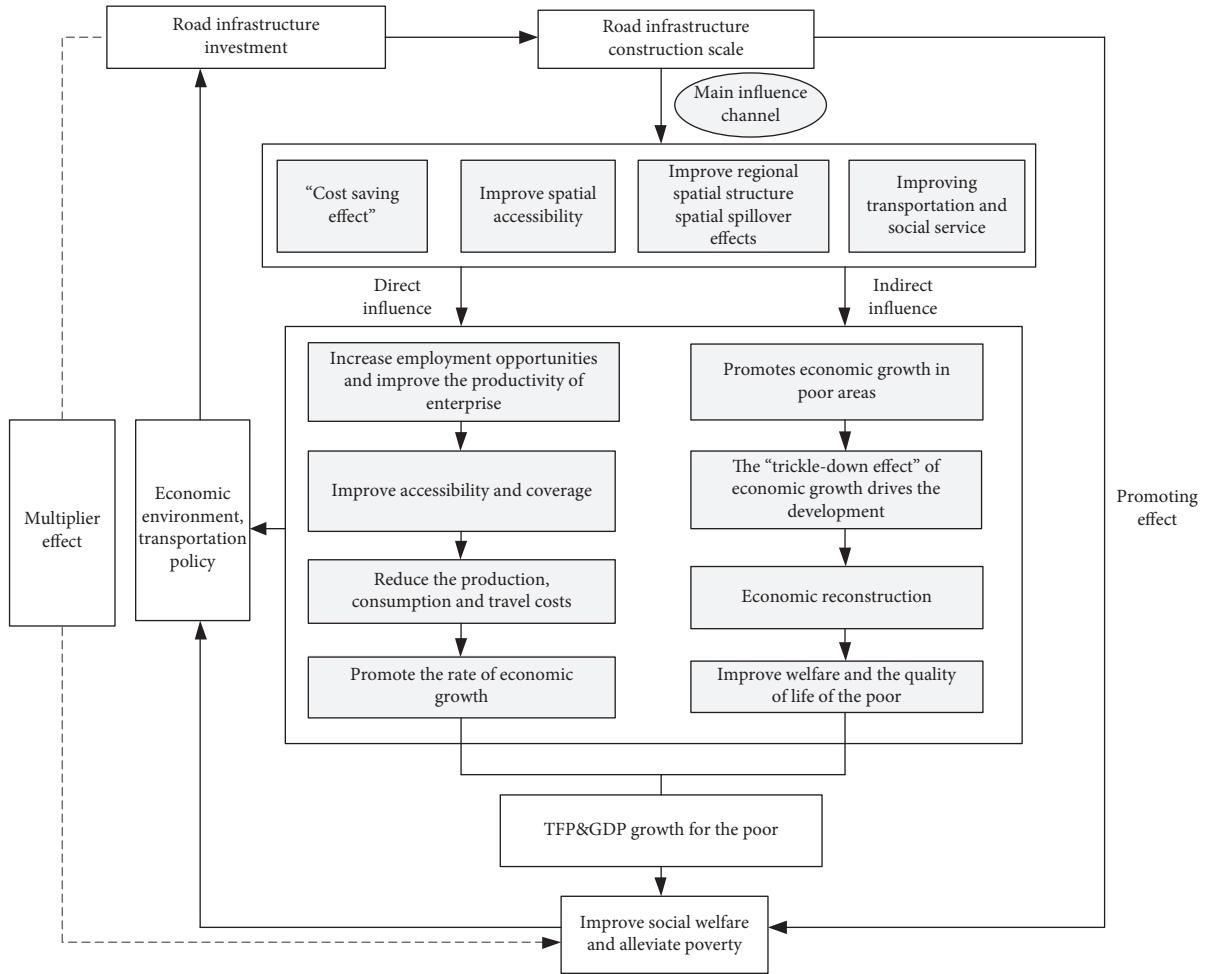
FIGURE 2: Theoretical framework for the influencing pathway between road infrastructure and local economy.

3. Materials and Methods

3.1. Data Sources. This paper takes 91 cities/counties in the concentrated contiguous destitute areas of Yunnan province as the research object. The research data are mainly composed of the basic geographic data and the social public service guarantee data as shown in Table 1.

3.2. Statistical Analysis and Modeling

3.2.1. Road Infrastructure Assessment Model. Before turning to theoretical and empirical investigation, we made some descriptive statistics on the road infrastructure and poverty incidence in the study area.

Yunnan province, as an underdeveloped region in the west, is located on the Yunnan–Guizhou plateau with complex terrain. Compared with other modes of transportation, the highway network has a large scale and is the leading mode of the comprehensive transportation system of the whole province. In this paper, the configuration of road infrastructure is used to represent the situation of road infrastructure in the whole province. Considering the practicability of the data, four indexes of highway mileage, highway mileage, highway proportion, and the number of passenger

stations above grade 2 are selected to construct the configuration status of road infrastructure at county level. In order to avoid the influence of subjective factors, the evaluation model of road infrastructure based on entropy weight method is constructed. The model expression is as follows:

$$z_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}},$$

$$e_j = -\frac{1}{\ln m} \sum_{i=1}^{m} z_{ij} \ln z_{ij},$$

$$d_j = 1 - e_j, \tag{1}$$

$$w_j = \frac{d_j}{\sum_{i=1}^{n} d_j},$$

$$\text{HIC}_i = \sum_{j=1}^{n} w_j z_{ij}.$$

In this formula, x_{ij} is the value of the j-th index of the i-th research unit; z_{ij} is the normalized value of the j-th index of the i-th research unit; m indicates the total number of research units; e_j is the entropy value of index j; d_j is the entropy redundancy of index j; w_j indicates the weight of index j; n is

TABLE 1: Data sources.

Dimension	Representative variables	Variable description	Data sources	Data source type
Basic geographic data	Vectorization of highway data	Vectorization of highway	Highway atlas 2015 (People's communication press)	Analytical data
	Per capita GDP	Per capita GDP	Yunnan Statistical Yearbook[1]	
	Annual net income per capita	The initial distribution and redistribution of income of urban households	China Statistical Yearbook For Regional Economy	
	Public budget revenue	Financial income for safeguarding and improving people's livelihood	China City Statistical Yearbook	
Socioeconomic factors	Total power of agricultural machinery	Reflecting the status of agricultural machinery owned by rural residents	Compilation of Transportation Statistics in Yunnan Province[2]	Statistical data
	Fixed telephone year-end user	Telephone penetration directly reflects the poverty level of regional residents	Statistical Bulletin on National Economy and Social Development in Yunnan Province	
	Urbanization rate	The ratio of the permanent population of the town to the total resident population of the district		
Social security	Fixed asset investment	Reproduction of fixed assets	Annual Statistical Bulletin of Counties and Districts	
	Number of social welfare adoption units	Reflecting the social welfare facilities and basic security status of the region		
	Per capita housing construction area	Average residential building area per person based on resident population		
Material capital	Aging proportion	The proportion of the elderly population to the total population		
	Rural road investment	Reflecting the state's investment in rural infrastructure		
Transportation facilities	County road pavement rate	Reflecting the quality and popularity of transportation infrastructure		
	Road network density			
	The population density	Population density is the number of people living on a unit of land		
Geographic capital	Grain planting area	Cultivated land is the most important material livelihood capital, and grain yield usually highly correlates with poverty		
	Annual average precipitation	Rainfall reflects regional ecological conditions		

[1]Yunnan Statistical Yearbook (1988-2017), Yunnan Provincial Bureau of Statistics (2000-2017), Yunnan Provincial. [2]Yunnan Province Traffic Statistics Compilation (2000-2017), Yunnan Provincial Department of Transportation (2000-2017).

the total number of evaluation indicators; and HIC_i is the road infrastructure grade score of the research unit i.

3.2.2. Coupling Harmonious Degree Model.

In order to obtain the level of coordination between road infrastructure and poverty rate at the county scale, a coupling coordination degree model was constructed. The model expression is as follows:

$$W_i = 2 \times \left[\frac{K_i \times H_i}{(K_i + H_i)^2} \right]^{(1/2)}, \qquad (2)$$

$$X_i = \sqrt{W_i \times Q_i}, \qquad (3)$$
$$Q_i = \alpha K_i + \beta H_i.$$

In this formula, W_i denotes the coupling degree of the area and county i; K_i and H_i denote the road

infrastructure grade score and poverty incidence rate of county i; X_i indicates the coupling coordination degree of the area and county i; Q_i indicates the comprehensive coordination index of the score of road infrastructure grade and poverty incidence in the area and county i; and α, β are the undetermined coefficients, which are generally 0.5.

3.2.3. Structural Equation Model.

The economic impact of transportation infrastructure investment is the most widely discussed issue, and various methodologies have been applied to address this problem. The most common approaches are the production function model and the cost function approach. [47–49]. Although they are widely used, there are still some problems. The production function model does not account for simultaneity or any noncontemporaneous effects among the different variables. Although the cost

function approach is able to show whether transportation investment can contribute to economic development and how this effect is realized, it also has its own shortcomings of spurious correlation, nonstationary, and noncointegration [46]. Compared with the above-mentioned approaches, structural equation model (SEM) is flexible and practical; it can be preperceived by correspondence canonical analysis. Therefore, SEM is employed to investigate the transport-region economic relationship, as well as the mediating factors. We perform the F test for variance homogeneity, the Shapiro–Wilk test for normality, and the standard deviation model for standardization. The maximum likelihood was employed to establish the SEMs for the concentrated contiguous destitute areas of Yunnan. To evaluate the model performance, we refer to several well established indicators, including the chi-square (χ^2) test, the goodness-of-fit index (GFI), the comparative fit index (CFI), the root mean squared error of approximation (RMSEA), and the adjusted goodness-of-fit index (AGFI). Model validity is evaluated by the critical ratio (CR) and item reliability (R^2). Acceptable values for the two indicators are recommended as |CR|>2.0 and $R^2 > 0.5$. Relationships among different variables are interpreted based on the standardized coefficients [9].

The initial structural equation model contains five latent variables and 30 observed variables; according to the test results, the structural equation model established in this paper is overidentified. The revised model includes 14 indicators reflecting the county's natural geographical conditions, resource endowments, traffic location, education level of residents, welfare protection, and social and economic development level. The model introduced potential variables on the basis of considering the path relationship between exogenous variables and endogenous variables.

$$y = By + \Gamma x + \zeta, \tag{4}$$

where y is the exogenous latent variables, represented by a vector composed of endogenous indicators; x is the endogenous latent variables, represented by a vector composed of exogenous indicators; B is the coefficient of interaction between endogenous latent variables; Γ is the effect coefficient of exogenous latent variables on endogenous latent variables (path coefficient); and ζ is the residual vector of the structural equation, reflecting the unexplained part of y in the equation.

The road infrastructure is based on structural equation model to construct poverty reduction path and the strength of research methods, mainly through path graph and effect value to analyze road infrastructure construction investment and how it affects the economic growth and poverty reduction, as well as how road infrastructure indirectly contributes to poverty reduction through other industries and capital, to characterize the spatial spillover effects of the road infrastructure and the effectiveness of comprehensive poverty alleviation.

4. Results

4.1. Analysis of Spatial Characteristics of Road Infrastructure and Poverty Incidence. Based on the data of highway infrastructure and poor population, formula (1) and poverty rate formula are used to calculate the road infrastructure grade score and poverty incidence in the research area, and ArcGIS spatial analysis technology is used to describe its spatial characteristics. The detailed analysis results are presented in Figures 3 and 4.

(1) Unbalanced distribution and unreasonable proportional structure of road network in Yunnan province were found. The rural road coverage rate is obviously insufficient, resulting in the low level of the overall road network.

(2) Yuanjiang valley and the south section of Yunling mountain, taken as the boundary, present a high-low distribution "transition zone" of road infrastructure from Deqin County to Hekou County, with east Yunnan and central Yunnan high-value areas in the east, low-value areas in the border area of Western Yunnan, and the high-value area and low-value area showing the characteristics of strong spatial agglomeration.

(3) The poverty in the study area presents a strong spatial heterogeneity. The areas with high poverty rate are mainly concentrated in the Diqing Tibetan area and the Wumeng Mountain area, forming "double cores" of different sizes.

4.2. Spatial Coupling Results of Road Infrastructure and Poverty Rates. On the basis of obtaining the characteristics of road infrastructure and poverty, formulas (2) and (3) are used to calculate the coupling and coordination value of road infrastructure and poverty rate in the district counties, and the coupling and coordination value is divided into five spatial relations according to the calculation results, as shown in Figure 5.

(1) About 70% of the district counties' road infrastructure and poverty rate are in a state of spatial coupling imbalance. Among them, as one of the "three districts and three states" in the deep poverty areas, Nujiang belongs to a typical type of maladjustment and common loss. Influenced by terrain, geological conditions, resource endowment, and other factors, the lag of road infrastructure and deep poverty are intertwined for a long time.

(2) The coordinated synchronous type is mainly distributed along the transportation network. For example, the continuous promotion of the construction of Kunming–Qiaojia expressway and Panzhihua-Zhaotong–Bijie–Zunyi railway will speed up the gradual improvement of the comprehensive transportation system of Qiaojia and Daguan Counties, and the steady improvement of economic development. The completion and opening of the high-speed railway between Yunnan and Guangxi have driven the rapid development of tourism, characteristic agriculture, and service industry in Maitreya, and the number of employees has increased significantly.

Road network grade
— National road
— Provincial road
— County road
— Township road

Grade of highway infrastructure
□ 4.23 ~ 26.38
■ 26.39 ~ 51.35
■ 51.36 ~ 102.57
■ 102.58 ~ 193.02
■ 193.03 ~ 305.68

(a)

(b)

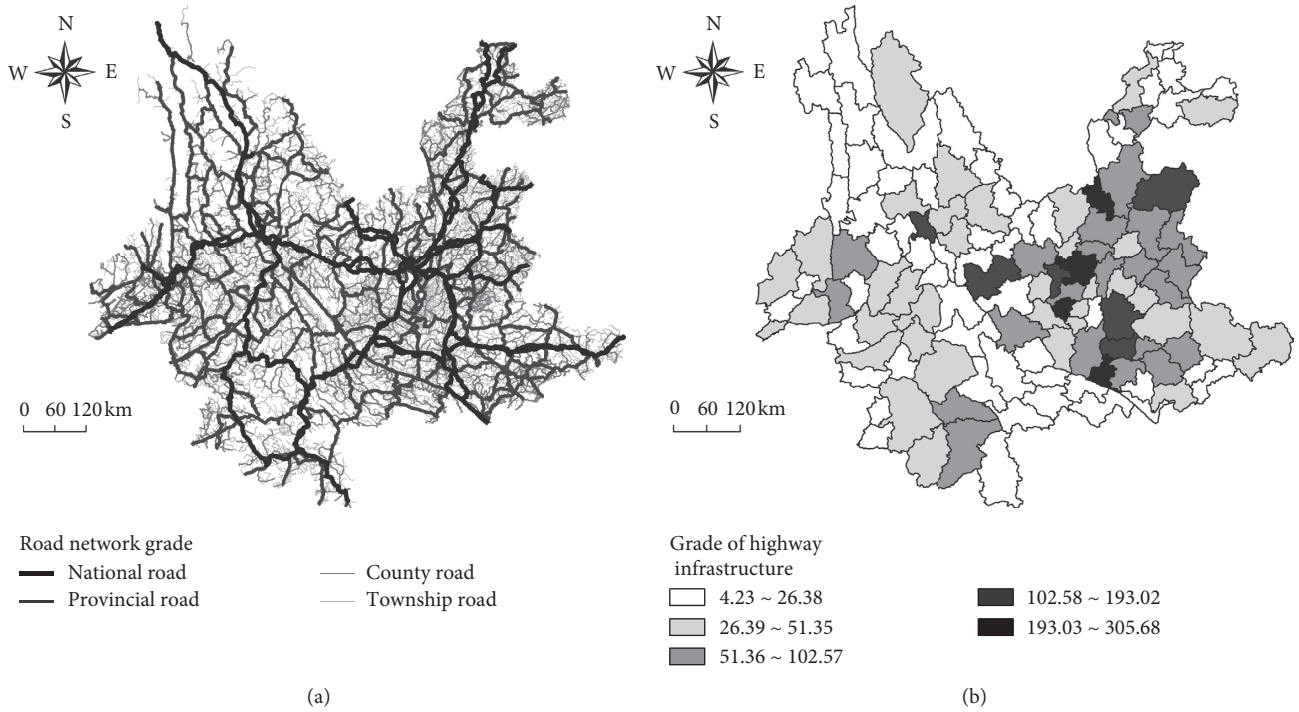FIGURE 3: Spatial distribution of road infrastructure network and hierarchy. (a) Road network grade. (b) Road infrastructure assessment.

Incidence of poverty in contiguous destitute areas
□ Non-district county
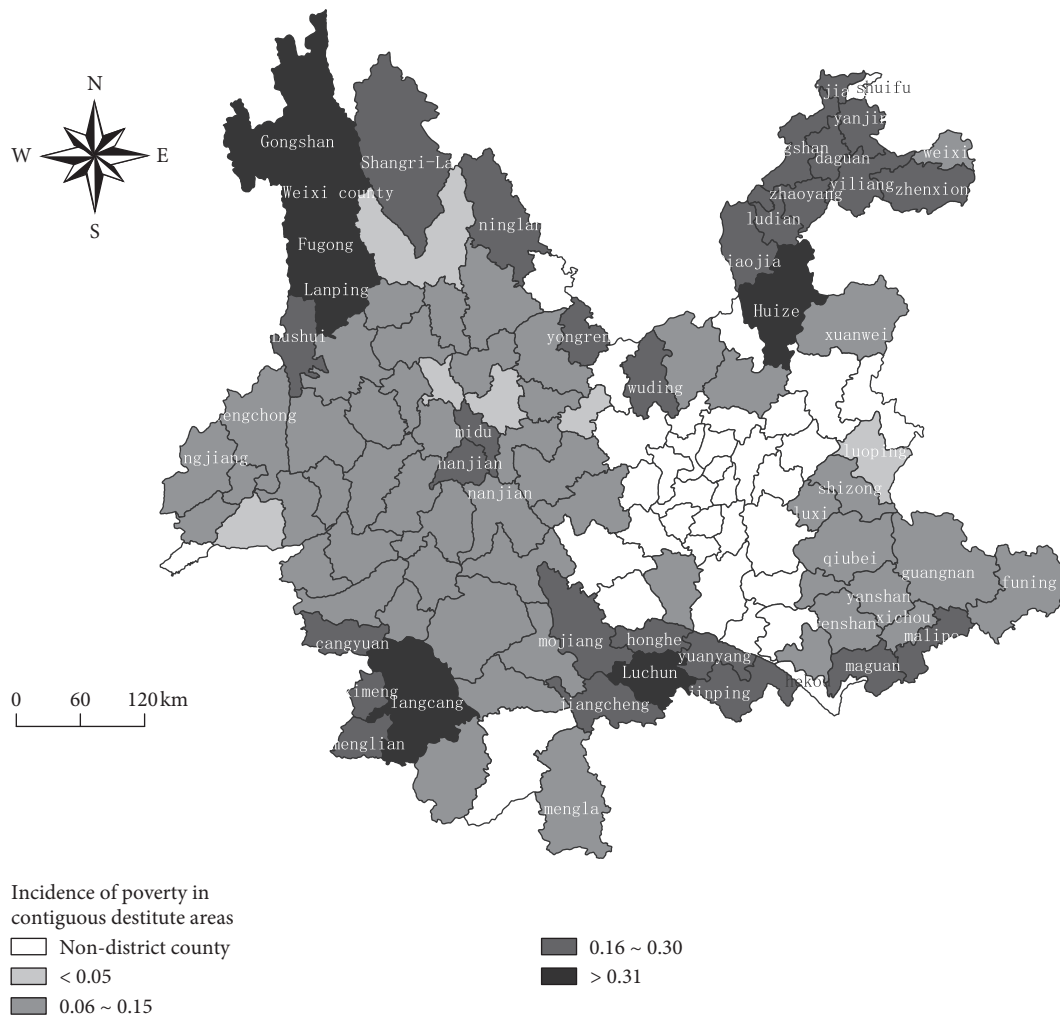■ < 0.05
■ 0.06 ~ 0.15
■ 0.16 ~ 0.30
■ > 0.31

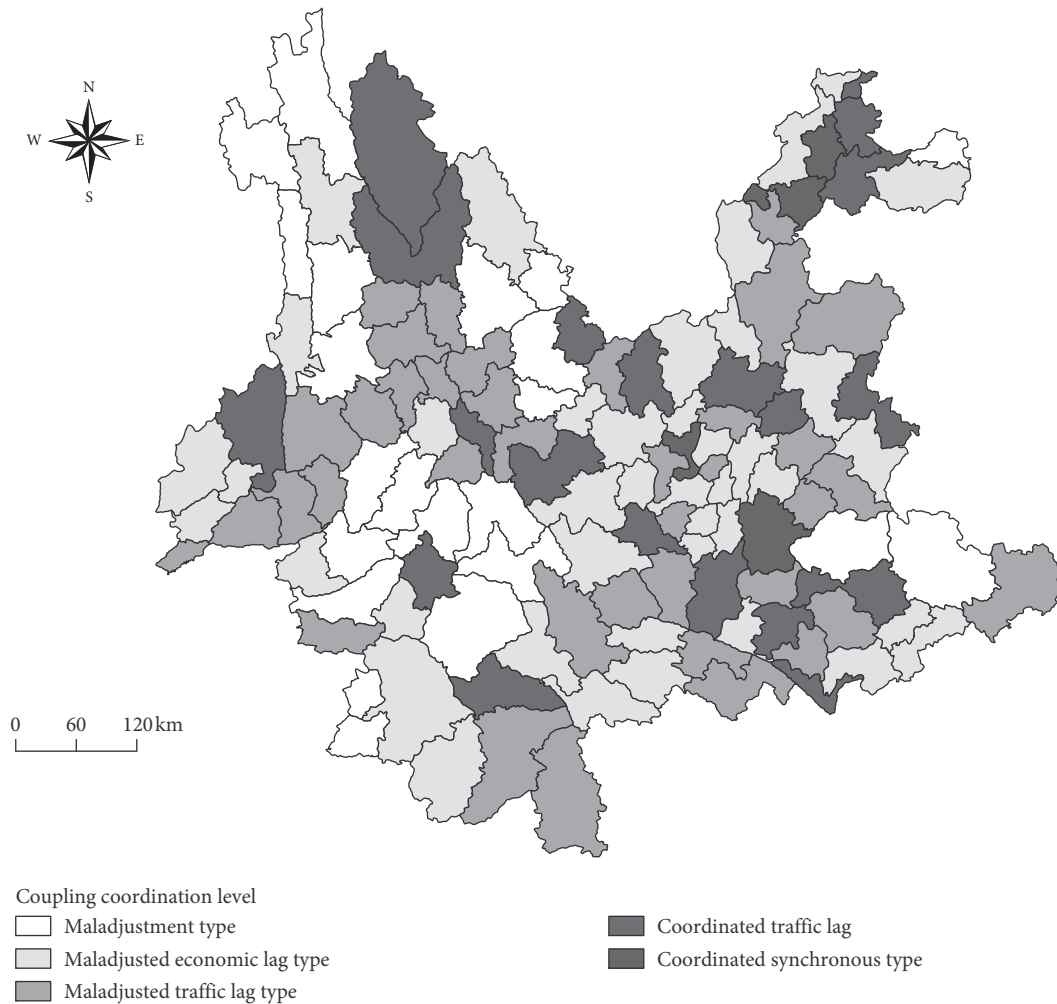FIGURE 4: Spatial distribution of poverty rate.

FIGURE 5: Spatial pattern of coupling coordination degree.

4.3. The Path and Intensity of Poverty Reduction on Road Infrastructure. On the basis of determining the coupling and coordination relationship between road infrastructure and poverty rate, the effect of poverty reduction becomes the focus of attention. Furthermore, structural equation model was used to quantify the path and intensity of action of road infrastructure investment and construction on poverty reduction.

The revised model overall fitness test and the optimal model parameter estimation results are shown in Tables 2 and 3. The analysis results show that the test results are generally matched, and the path coefficients in the model basically reach the significance of 1% level. The following path coefficients are acceptable, and the model variance estimates are all positive and consistent with the level of significance.

In the structural equation model, the path coefficient value reflects the impact of different variables such as road infrastructure, social capital, and geographical capital on poverty reduction. Investment and construction of road infrastructure stimulate the economy to achieve the ultimate goal of reducing poverty through direct or indirect diversification of routes, as shown in Figure 6.

(1) In terms of the direct poverty reduction effect, the road infrastructure investment and construction can directly affect the poverty reduction, with a path coefficient of 0.105, which is very significant. However, the direct effect is far lower than the indirect poverty reduction effect (0.830) brought by the road infrastructure investment and construction, which affects geographical location, fixed asset investment, secondary and tertiary industries development, and urbanization. This shows that the contribution of transport infrastructure investment and construction to poverty reduction is more indirect in the form of spatial spillover, network attributes, and externality of road infrastructure construction, which play an important role in economic development and poverty reduction. The enlargement of road networks would definitely accelerate capital and resource flow, as well as human movement. Consequently, the local interactions with outside economy are increased. Such interactions are considered critical drivers for economic growth in less developed areas. In particular, inputs from other sectors can be created and increased by road network

TABLE 2: Evaluation results of the modified model.

Statistical test	X^2	df	X^2/df	RMSEA	CFI	NFI	GFI	IFI	AGFI
Fitness threshold	—	—	1 < NC < 3, good 3 < NC < 5, not too bad	RMSEA < 0.5, good 0.5 < NC < 0.8, not too bad	>0.9	>0.9	>0.9	>0.9	>0.9
Inspection result	212.584	75	2.826	0.484	0.961	0.934	0.968	0.985	0.944
Adaptation judgment	—	—	Good	Good	Perfect	Perfect	Perfect	Perfect	Perfect

TABLE 3: Estimation results of the variable coefficient and significance test.

Acting path	Estimate	SE	CR	P	Label	Standardized estimates
Material capital < --- road infrastructure	0.470	0.270	2.742	0.002	par_21	0.323
Geographical capital < --- road infrastructure	0.020	0.002	8.843	***	par_29	0.669
Social capital < --- material capital	4.697	2.485	1.980	0.039	par_13	0.458
Social capital < --- geographical capital	385.159	64.215	5.998	***	par_14	0.757
Traffic dominance < --- road infrastructure	1.000					0.804
Highway network density < --- road infrastructure	0.146	0.030	4.818	***	par_1	0.572
Poverty reduction effect < --- material capital	3599.722	3506.741	3.027	***	par_18	0.263
Poverty reduction effect < --- social capital	1509.292	576.523	2.618	**	par_19	1.132
Poverty reduction effect < --- geographical capital	7535.539	224994.298	3.033	***	par_25	0.011
Total power of agricultural machinery < -- social capital	1.000					0.702
Revenue from the public budget < ---highway network density	2291.362	534.740	4.285	***	par_17	0.166
Minority share < ---transport advantage	1414.846	263.234	5.375	***	par_27	0.284
Rural road investment < --- road infrastructure	0.053	0.017	3.201	***	par_2	0.230
Per capita housing construction area < --- material capital	1.000					0.163
Number of social welfare adoption units < --- material capital	0.755	0.451	2.675	**	par_3	0.213
The proportion of aging < --- material capital	0.006	0.003	2.986	***	par_4	0.412
Grain production < --- geographical capital	1558.083	1032.655	2.809	***	par_5	0.131
Urbanization rate < ---poverty reduction effect	0.007	0.001	7.897	***	par_6	0.615
Fixed-line users at the end of the year < --- social capital	699.746	114.571	6.108	***	par_7	0.401
Per capita GDP < ---poverty reduction effect	1.000					0.702
Revenue from the public budget < --- social capital	0.000	0.000	20.600	***	par_9	1.109
Fixed assets investment < --- social capital	8.685	0.681	12.760	***	par_20	0.959
Population density < --- geographical capital	1.000					0.726
The proportion of aging < --- total power of agricultural machinery	0.000	0.000	6.899	***	par_22	0.449
Per capita housing construction area < --- proportion of ethnic minorities	0.001	0.000	11.120	***	par_28	0.627

"<---" indicates the orientation in which the utility occurred; **$P < 0.01$, ***$P < 0.001$.

construction, and the multiplier economic effect is subsequently stimulated. This finding suggests that the realization of the effect of road infrastructure on economic growth needs the synergy of institutional environment improvement, human capital improvement, science and education technology level improvement, and other factors. When analyzing the causes of economic growth, multidimensional factors need to be taken into account in order to improve the social welfare and reduce transport poverty in poverty-stricken areas. This discovery is in line with the past literature reporting that the spatial structure evolution was closely linked with road infrastructure construction [50, 51].

(2) In terms of the indirect poverty reduction effect, there is a significant difference in the path and degree of the road infrastructure to promote economic growth and ultimately reduce poverty. This paper mainly works through five poverty reduction path modes of road infrastructure. Among them, " road infrastructure ⟶ social capital ⟶ poverty reduction effect" has the highest path coefficient (0.740), and the contribution of fixed asset investment and public budget income to social capital is 0.959 and 1.109, respectively, which are higher than other poverty reduction factors.

(3) Road infrastructure investment and construction can guide economic development and reduce poverty by acting on geographic capital, material capital and social capital. The overall action coefficient is 0.740, which is higher than the poverty reduction effect of the two paths of "road infrastructure ⟶ geographic capital/material capital ⟶ poverty reduction effect". The path coefficient of "road infrastructure ⟶ geographical capital ⟶ social capital ⟶ poverty reduction effect" is 0.573, which is higher than that of "road infrastructure ⟶ material capital ⟶ social capital ⟶ poverty reduction effect" (0.167).
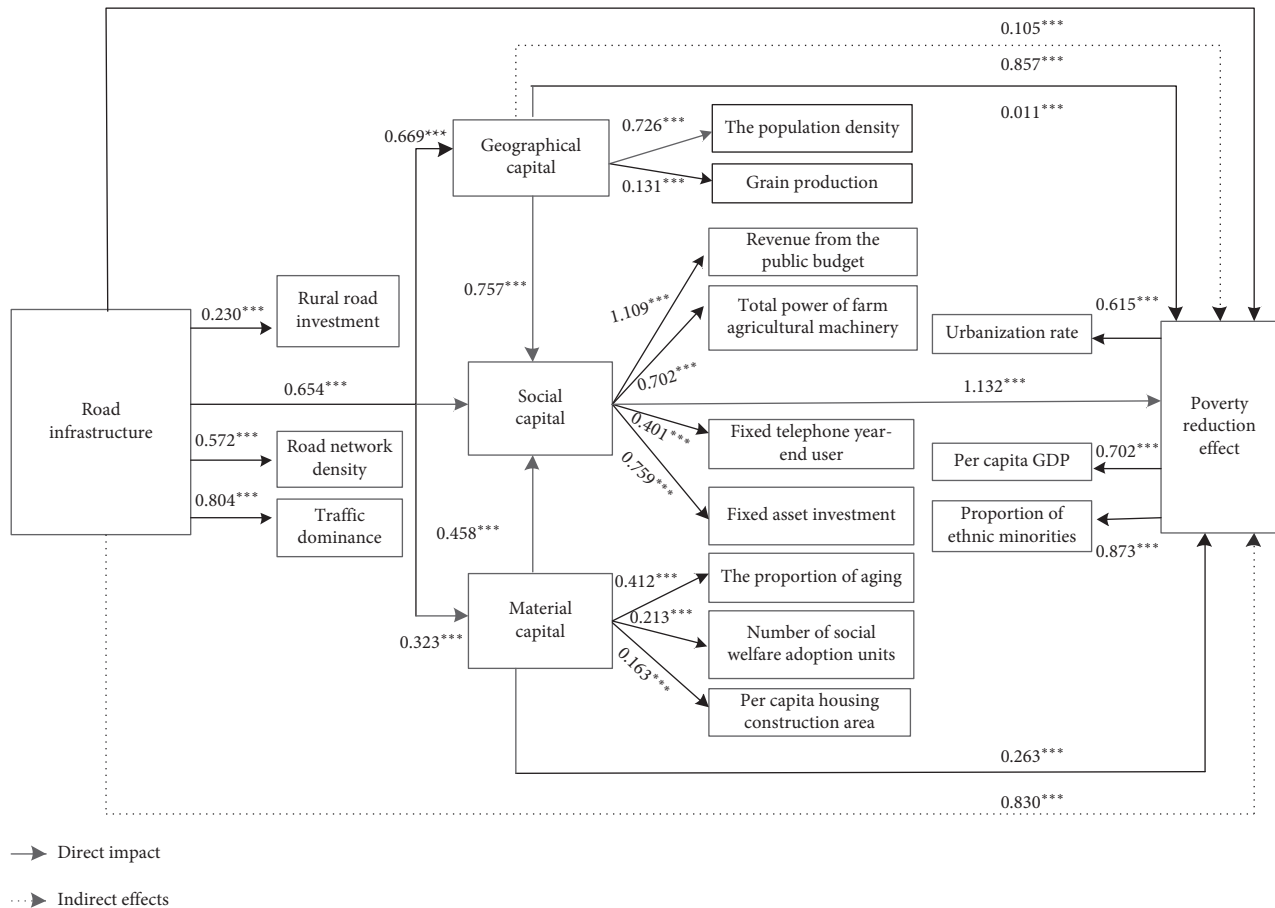
Figure 6: Estimation results of the SEM for poverty alleviation path of road infrastructure.

(4) Either the investment and construction of transportation infrastructure itself or the associated capital and industrial development brought by it are obviously conducive to economic growth and employment of the poor. Taking the construction of highway transportation investment as an example, in the link of road production and construction, highway transportation investment and construction can attract a large amount of labor to increase nonagricultural employment opportunities, reduce transportation costs, and improve production efficiency. In the link of traffic operation, highway traffic investment and construction can promote the location choice, productivity distribution, and production factor agglomeration of enterprises.

5. Conclusion and Discussion

5.1. Conclusion. In this paper, the evaluation model of transportation infrastructure based on entropy weight method and a coupling coordination degree model are established, obtaining the spatial characteristics and spatial coupling results of road infrastructure and poverty incidence. On that basis, a structural equation model has been established to describe the relationships between the variables, and the relationship between transportation

investment and regional economic growth is estimated. Based on the results of empirical study, the following key conclusions are drawn.

(1) The structural equation model confirms the significant effect of road infrastructure investment and construction on economic growth and poverty reduction, but the path and degree of poverty reduction are quite different. The direct effect of road infrastructure investment and construction on poverty reduction is far less than the indirect effect, and the diversification path is produced indirectly through the form of spatial spillover. Among the five path models of poverty reduction effects on road infrastructure, "road infrastructure ⟶ social capital ⟶ poverty reduction effect" has the highest path coefficient of poverty reduction, and the effect of poverty reduction is the best.

(2) The contribution rate of the number of passenger stations above grade 2 and county road pavement rate is higher than those of other traffic indicators, which is the most important poverty reduction factor in the study area.

(3) About 70% of the district counties' road infrastructure and poverty rate are in a state of spatial coupling imbalance. The coordinated synchronous

type is mainly distributed along the transportation network.

The research results can provide theoretical and practical basis for the formulation of transportation poverty alleviation strategy, so as to effectively play the leading and supporting role of transportation infrastructure in poverty alleviation and rural revitalization. In order to more comprehensively analyze the impact of transportation supply on economic growth and poverty reduction, based on the theory of spatial equity and the principle of regional deprivation compensation, we put forward the poverty alleviation and optimization strategy of traffic resource allocation in the concentrated contiguous destitute areas from the perspective of spatial equity:

(1) Speeding up the building of a transportation network that connects the poverty-stricken areas with the outside world and the inside world and connects the villages and townships

(2) Innovating ways of providing transportation to help the poor and advancing a new comprehensive approach of transportation poverty alleviation

(3) Developing public-private partnerships (PPP) in various forms and innovating investment and financing mechanisms

Considering this strategic decision-making in the statistical model could produce interesting and useful policy insights.

5.2. Discussion. Transportation infrastructure has complex interactions with regional economic systems. Prior studies indicate that impacts of transportation infrastructure are more pronounced in developing countries, as the prevalence of infrastructure is lower in these countries [52]. This paper builds several models to uncover the contributing factors in mediating the relationships between the road infrastructure and local economy in the concentrated contiguous destitute areas of Yunnan. We found that the established models are efficient in revealing the relationship between the road infrastructure and the less developed places' economy. Previous studies advocate road infrastructure as a primary economic catalyst in less developed areas, as it can increase the accessibility to market, eliminate constraints to social services, and provide employment [6, 7]. The results for the concentrated contiguous destitute areas of Yunnan are in accordance with earlier findings.

This paper demonstrates that SEM is efficient in our case. However, SEM is not always strong in proving causality. Regional and local policies regarding road infrastructure investment promoting economic growth differ significantly across various regions in different countries; when applying the demonstrated methodology to other areas worldwide, it requires caution and further verification. Though our study advances the understanding of the relationships between the road infrastructure and regional economic, it still has certain shortcomings. First, all the indicators of the article are chosen based on empirical literature review and data availability; subjectivity is unavoidably associated with the selection methods by which the exploratory variables can be selected more scientifically. Second, the SEMs are established by using the maximum likelihood approach in this study. However, whether or not this approach is a methodological advantage remains unproven. Third, this paper examines the concentrated contiguous destitute area at the provincial level; to deepen the understanding of the relationship between the road infrastructure and local economy, future studies should investigate and compare different regions with different levels of development. Fourth, road infrastructure is not the only type of road infrastructure. Since the highway network has a large scale and is the leading mode of the comprehensive transportation system of the Yunnan province, our study does not consider railways and high-speed rails. Some studies have evidenced the effect of high-speed rails on the economic characteristics; consequently, the exclusion of railways may overestimate the empirical findings in this study. Further studies should control the railways and high-speed rails, to isolate the individual effect of road infrastructure on the urban economy.

Data Availability

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

Disclosure

The funders had no role in the design of the study or in the collection and interpretation of the data.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] R. Kanbur, "Economic policy, distribution and poverty: the nature of disagreements," *World Development*, vol. 29, no. 6, pp. 1083–1094, 2001.

[2] D. Narayan, R. Patel, K. Schafft et al., *Voices of the Poor Can Anyone Hear Us?*, Oxford University Press for the World Bank, New York, NY, USA, 2000.

[3] R. Vakis, J. Rigolini, and L. Lucchetti, *Left behind: Chronic Poverty in Latin America and the Caribbean*, World Bank publication, Washington, DC, USA, 2016.

[4] ADB (Asian Development Bank), *Sustainable Transport Initiative: Operational Plan*, ADB, Philippines, PA, USA, 2010.

[5] Y. Liu, Z. Liu, and R. Jia, "DeepPF: a deep learning based architecture for metro passenger flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 18–34, 2019.

[6] R. Asomani-Boateng, R. J. Fricano, and F. Adarkwa, "Assessing the socio-economic impacts of rural road improvements in ghana: a case study of transport sector program support (II)," *Case Studies on Transport Policy*, vol. 3, no. 4, pp. 355–366, 2015.

[7] M. Iacono and D. Levinson, "Mutual causality in road network growth and economic development," *Transport Policy*, vol. 45, pp. 209–217, 2015.

[8] S. E. Polzin, "Transportation/land-use relationship: public transit's impact on land use," *Journal of Urban Planning and Development*, vol. 125, no. 4, pp. 135–151, 1999.

[9] H. You, "Agricultural landscape dynamics in response to economic transition: comparisons between different spatial planning zones in Ningbo region, China," *Land Use Policy*, vol. 61, pp. 316–328, 2017.

[10] S. Su, Q. Zhang, J. Pi, C. Wan, and M. Weng, "Public health in linkage to land use: theoretical framework, empirical evidence, and critical implications for reconnecting health promotion to land use policy," *Land Use Policy*, vol. 57, pp. 605–618, 2016.

[11] M. Wachs and T. G. Kumagai, "Physical accessibility as a social indicator," *Socio-Economic Planning Sciences*, vol. 7, no. 5, pp. 437–456, 1973.

[12] S. Hanson and P. Hanson, "Gender and urban activity patterns in uppsala, Sweden," *Geographical Review*, vol. 70, no. 3, pp. 291–299, 1980.

[13] A. Armstrong-Wrigh, *Urban Transit Systems*, World Bank Technical Papers, Washington, DC, USA, 1986.

[14] D. A. Aschauer, "Is public expenditure productive?" *Journal of Monetary Economics*, vol. 23, no. 2, pp. 177–200, 1989.

[15] D. A. Aschauer, "Does public capital crowd out private capital?" *Journal of Monetary Economics*, vol. 24, no. 2, pp. 171–188, 1989.

[16] D. A. Aschauer, "Public investment and productivity growth in the group of seven," *Journal of Monetary Economics*, vol. 13, no. 5, pp. 17–25, 1989.

[17] C. A. Gannon and Z. Liu, *Poverty and Transport*, World Bank Publications, Washington, DC, USA, 1997.

[18] M. G. Boarnet, "Spillovers and the locational effects of public infrastructure," *Journal of Regional Science*, vol. 38, no. 3, pp. 381–400, 1998.

[19] G. Porter, "Transport services and their impact on poverty and growth in rural sub-saharan africa: a review of recent research and future research needs," *Transport Reviews*, vol. 34, no. 1, pp. 25–45, 2014.

[20] R. Benevenuto and B. Caulfield, "Poverty and transport in the global south: an overview," *Transport Policy*, vol. 79, pp. 115–124, 2019.

[21] A. Church, M. Frost, and K. Sullivan, "Transport and social exclusion in London," *Transport Policy*, vol. 7, no. 3, pp. 195–205, 2000.

[22] J. Farrington and C. Farrington, "Rural accessibility, social inclusion and social justice: towards conceptualisation," *Journal of Transport Geography*, vol. 13, no. 1, pp. 1–12, 2005.

[23] N. Cass, E. Shove, and J. Urry, "Social exclusion, mobility and access," *The Sociological Review*, vol. 53, no. 3, pp. 539–555, 2005.

[24] N. R. Velaga, M. Beecroft, J. D. Nelson, D. Corsar, and P. Edwards, "Transport poverty meets the digital divide: accessibility and connectivity in rural communities," *Journal of Transport Geography*, vol. 21, pp. 102–112, 2012.

[25] K. Lucas, "Transport and social exclusion: where are we now?" *Transport Policy*, vol. 20, pp. 105–113, 2012.

[26] ITF (International Transport Forum), "Income inequality, social inclusion and mobility," *Roundtable Report*, vol. 164, p. 200, 2017.

[27] K. Lucas, G. Mattioli, E. Verlinghieri, and A. Guzman, "Transport poverty and its adverse social consequences," *Proceedings of the Institution of Civil Engineers - Transport*, vol. 169, no. 6, pp. 353–365, 2016.

[28] T. Litman, "Measuring transportation: traffic, mobility and accessibility," *Institute of transportation engineers*, vol. 73, no. 10, p. 28, 2003.

[29] Z. Liu, S. Wang, K. Huang, J. Chen, and Y. Fu, "Practical taxi sharing schemes at large transport terminals," *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, pp. 596–616, 2019.

[30] X. Qu, Y. Yu, M. Zhou et al., "Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: a reinforcement learning based approach," *Applied Energy*, vol. 257, pp. 1–11, 2020.

[31] J. Luo, L. Tian, L. Luo, H. Yi, and F. Wang, "Two-step optimization for spatial accessibility improvement: a case study of health care planning in rural China," *BioMed Research International*, vol. 2017, Article ID 2094654, 12 pages, 2017.

[32] R. C. Nesbitt, S. Gabrysch, A. Laub et al., "Methods to measure potential spatial access to delivery care in low- and middle-income countries: a case study in rural Ghana," *International Journal of Health Geographics*, vol. 13, no. 1, p. 25, 2014.

[33] D. J. Weiss, A. Nelson, H. S. Gibson et al., "A global map of travel time to cities to assess inequalities in accessibility in 2015," *Nature*, vol. 553, no. 7688, pp. 333–336, 2018.

[34] A. Iimi, A. K. F. Ahmed, E. C. Anderson et al., *New Rural Access Index: Main Determinants And Correlation To Poverty (English)*, World Bank Publication, Washington, DC, USA, 2016.

[35] A. K. Sarkar and D. Ghosh, "Integrated rural accessibility planning (IRAP)," *Progress in Development Studies*, vol. 8, no. 3, pp. 241–259, 2008.

[36] R. H. M. Pereira, T. Schwanen, and D. Banister, "Distributive justice and equity in transportation," *Transport Reviews*, vol. 37, no. 2, pp. 170–191, 2017.

[37] P. Jiwattanakulpaisarn, R. B. Noland, and D. J. Graham, "Causal linkages between highways and sector-level employment," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 4, pp. 265–280, 2010.

[38] N. Yu, G. de Roo, M. de Jong, and S. Storm, "Does the expansion of a motorway network lead to economic agglomeration? evidence from China," *Transport Policy*, vol. 45, pp. 218–227, 2016.

[39] G. M. Ahlfeldt, K. Moeller, and N. Wendland, "Chicken or egg? The PVAR econometrics of transportation," *Journal of Economic Geography*, vol. 15, no. 6, pp. 1169–1193, 2015.

[40] S. Bougheas, P. O. Demetriades, T. P. Mamuneas et al., "Infrastructure, specialization, and economic growth," *Canadian Journal of Economics/Revue Canadienne d'Economique*, vol. 33, no. 2, pp. 506–522, 2000.

[41] M. A. Beyzatlar, M. Karacal, and H. Yetkiner, "Granger-causality between transportation and GDP: a panel data approach," *Transportation Research Part A: Policy and Practice*, vol. 63, pp. 43–55, 2014.

[42] G. Chi, "The impacts of highway expansion on population change: an integrated spatial approach," *Rural Sociology*, vol. 75, no. 1, pp. 58–89, 2010.

[43] J. Guo, J. Guo, and J. Xia, "Econometrical investigation on infrastructure investment and economic development in China: a case study using vector autoregression approach," *KSCE Journal of Civil Engineering*, vol. 15, no. 3, pp. 561–567, 2011.

[44] N. Yu, M. De Jong, S. Storm, and J. Mi, "Transport infrastructure, spatial clusters and regional economic growth in China," *Transport Reviews*, vol. 32, no. 1, pp. 3–28, 2012.

[45] International Fund for Agricultural Development, *Rural Poverty Report*, International Fund for Agricultural Development, Rome, Italy, 2011.

[46] X. Jiang, X. He, L. Zhang, H. Qin, and F. Shao, "Multimodal transportation infrastructure investment and regional economic development: a structural equation modeling empirical analysis in China from 1986 to 2011," *Transport Policy*, vol. 54, no. 2, pp. 43–52, 2017.

[47] J. L. Fan and B. L. Bai, "Impact on economic growth of infrastructure capital: estimate using CES production function method," *Economic*, vol. 11, pp. 10–13, 2004.

[48] W. Qi, Y. Wang, Y. Bie, and J. Ren, "Prediction model for bus inter-stop travel time considering the impacts of signalized intersections," *Transportmetrica A: Transport Science*, vol. 02, pp. 1–19, 2020.

[49] Y. Bie, X. Xiong, Y. Yan, and X. Qu, "Dynamic headway control for high-frequency bus line based on speed guidance and intersection signal adjustment," *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 1, pp. 4–25, 2020.

[50] T. T. Ting and Y. Edward, "Transportation and economic growth in China: a heterogeneous panel cointegration and causality analysis," *Journal of Transportation Geography*, vol. 73, pp. 120–130, 2018.

[51] H. Li, Y. Liu, and K. Peng, "Characterizing the relationship between road infrastructure and local economy using structural equation modeling," *Transport Policy*, vol. 61, pp. 17–25, 2018.

[52] D. Banister and J. Berechman, *Transport Investment and Economic Development*, UCL Press, London, UK, 2000.

Research Article

Characteristics Analysis of Bus Stop Failure Using Automatic Vehicle Location Data

Rui Li,[1] Xin Xue,[1] and Hua Wang [ID][2]

[1]*College of Civil and Transportation Engineering, Hohai University, 1 Xi Kang Road, Nanjing 210098, China*
[2]*Department of Civil & Environmental Engineering, National University of Singapore, 21 Lower Kent Ridge Road, 117576, Singapore*

Correspondence should be addressed to Hua Wang; hwang191901@gmail.com

Queue forming behind a bus stop on an urban street is common and a traffic bottleneck usually occurs around the bus stop area. The bus stop failure means arriving buses cannot move into the bus stop due to limited capacity but have to wait for available loading areas. It is related with the transit operation level. Traditionally, the failure rate (FR), defined as the percentage of buses that arrives at the bus stop to find all loading areas occupied, is adopted in bus capacity analysis. However, the concept of FR is unable to quantitatively analyze failure characteristics in terms of its dispersion and uncertainty over time. Therefore, in this paper, we propose a new index called failure duration rate (FDR) to evaluate the bus stop failure, which can characterize waiting time for traffic delay calculation and capacity drop estimation. The automatic vehicle location data at eight bus stops in Wujiang District Suzhou, China, over 56 working days, are used to analyze the temporal characteristics of FR and FDR. We next examined the failed service duration characteristics during peak hours at the eight bus stops. Based on these characteristics analyses, we then proposed a Distribution Fitting and Cumulative Distribution Correlation (DF-CDC) approach to explore the correlation between FDR and FR at the same cumulative distribution function levels and validated the bus stop failure performance using the cross-validation method. The analysis results revealed that (i) FR fluctuates more significant than FDR, (ii) FDR is a more robust index than FR in describing the traffic characteristics incurred by bus stop failures, and (iii) FDR performs better in failure characteristics analysis than FR.

1. Introduction

A bus stop serving a large number of bus lines can experience a condition known as bus stop failure due to limited capacity and high passenger demand, which will negatively affect the punctuality and reliability of bus service and also bring about delays to other traffic. The more frequently the bus stop failure takes place, the lower the transit system level of service (LOS) is [1]. However, irregular traffic flow characteristics concerned with bus stop failure are difficult to be captured and quantified because of its dispersion and uncertainty over time [2]. Without a doubt, bus stop failure will significantly affect traffic characteristics at bus berths and adjacent lanes. Typically, the failure rate is proposed for analyzing and evaluating the influence of bus stop failure at bus berths.

Bus stop constitutes one potential bottleneck to interrupt traffic flow, which will deteriorate the level of transit operational service [3]. Transit operation parameters of buses served at loading areas, including dwell time [4, 5], headway [6, 7], capacity [8, 9], queue length [10, 11], and bunching characteristics [12, 13], are analyzed for evaluating the impact of bus stop failures. The failed service will increase bus waiting time for passenger boarding and alighting, and its impact can be measured by an index called failure rate (FR), which is defined as the percentage of buses that arrive at stops to find all berths are full [1]. Wang et al. analyzed the correlation between failure rate and four kinds of transit dwell and arrival characteristics and proposed a diffusion approximation method [14]. As one of the desired level indexes of transit operation, the FR could assess the change of capacity and LOS of bus berths. Failure probability and

dwell time variability can develop the function of bus queue length, which can better reflect the effects on bus stop capacity [15]. A parameter "Z" associating with the desired failure rate (under the assumption of standard normal distribution fitting) is developed to account for the fluctuation in bus dwell time in bus loading areas, and the design failure rates for urban and rural areas are recommended for estimating bus berth capacity [1]. The failure rate of curbside bus stops can be influenced by a serial of factors (such as bus arrival distribution type, bus arrival rate, bus berth maximal service rate, and bus service time variation), and the normalized capacity and incremental change (for multiberth stops) in capacity at different failure rate levels are proposed [8]. Moreover, the analysis process of failure rate is a poor proxy and suggests choosing the average waiting delay for evaluating the bus berth LOS is expressed comprehensively [16].

The data associated with vehicle location are utilized for analyzing and predicting traffic flow characteristics, which can be effectively applied to transit operational characteristics analysis [17–20], bus schedule optimization design [21, 22], bus lane planning and control strategy [23–25], and transport network flow estimation [26–28]. Several findings of characteristics analysis for public transit LOS at bus stop appear in the relevant literature. A data platform for monitoring patterns of bus operation is developed, which is primarily composed of data acquired from the ITS system in Beijing, China. A multilevel framework for transit performance analysis is proposed considering several transit operational factors [29]. Based on automated vehicle location (AVL) data, lots of statistical parameters about travel times are analyzed for evaluating the performance of bus routes with transit priority facilities, and these tests indicate spatial and temporal characteristics are the most potent feature [30]. A regression method (using LS-SVM) is developed for exploring bunching patterns of buses halting at the stop area, and the headway irregularity pattern is analyzed using transit smart card data [31]. A probabilistic method considering the interference between buses, using the loading areas, is established for predicting travel time of buses using trajectory and ID card data, which can reflect buses' dwell time distribution pattern well [32].

Although several findings of bus stop failure analysis using the failure rate appear in the relevant literature, the duration time of bus stop failure is rarely mentioned. Additionally, little research has been observed using AVL data to analyze the characteristics of bus stop failure. In this paper, failure duration time is utilized for evaluating bus stop failure, and a measure called failure duration rate is proposed for failure analysis utilizing collected transit automated vehicle location (AVL) data. The characteristics analysis of bus stop failure using AVL data can provide valuable information for transit operation optimization to the public transit authority.

The remainder of this research is organized as follows. In Section 2, a characteristic index called Failure duration rate (FDR) is developed for bus stop failure characteristic evaluation. Section 3 explores the failed duration characteristics at different failure levels based on the AVL data

collected from eight bus stops in Wujiang District of Suzhou in China. In Section 4, a correlation analysis between FDR and FR is carried out by using a "Distribution Fitting and Cumulative Distribution Correlation (DF-CDC)" analysis approach. Section 5 concludes the paper.

2. Characteristic Analysis Indexes

2.1. Failure Rate. A bus stop failure occurs when a bus arrives at the loading areas but with no available berth to use. The failure rate (FR), defined as the percentage of bus queuing for moving into bus berths occupied by other dwelling buses [14], could be formulated by

$$FR = \Pr(n > b), \tag{1}$$

where n is the number of buses halted at a bus stop and b is the number of bus berths.

In general, transit vehicles' dock at a bus stop (including the served and waiting buses) will obey the first-in-first-out rule and usually disperse in an independent manner to each other. The probability for the case without adequate berths at a bus stop could be calculated by

$$\Pr(n > b) = 1 - \Pr(n = 1) - \Pr(n = 2) - \cdots - \Pr(n = b). \tag{2}$$

When the berths at a bus stop are all occupied by buses for passenger boarding and alighting, the number of active buses served at the bus stop is equal to the number of berths ($n = b$). And the probability of this kind of bus arrival can be approximated by the value of $Q_s^{n=b}$ divided by Q_s:

$$\Pr(n = b) \approx \frac{Q_s^{n=b}}{Q_s}, \tag{3}$$

where $Q_s^{n=b}$ is the occurrence times for ($n = b$) during a given observation period and Q_s is the total number of arriving buses during the same time duration.

Then, we have

$$\Pr(n > b) \approx 1 - \frac{Q_s^{n=1}}{Q_s} - \frac{Q_s^{n=2}}{Q_s} - \cdots - \frac{Q_s^{n=b}}{Q_s}$$
$$= \frac{Q_s - (Q_s^{n=1} + Q_s^{n=2} + \cdots + Q_s^{n=b})}{Q_s}. \tag{4}$$

To facilitate our presentation, we denote the right hand of equation (4) by

$$R_f = \frac{Q_s - (Q_s^{n=1} + Q_s^{n=2} + \cdots + Q_s^{n=b})}{Q_s}. \tag{5}$$

In other words, R_f stands for the FR of a bus stop, which reflects the level of failure (LOF) for bus loading areas.

2.2. Failure Duration Rate. As the berths of a bus stop are occupied, the next arriving bus needs to queue at street lanes and exerts negative impacts on blocking movements of other vehicles along the same street lane. As a result, traffic delay goes up, and travel time reliability would be reduced [1]. It is an important and challenging task to analyze these adverse

effects quantitatively. In general, the longer the bus stop failure service lasts, the worse the traffic efficiency evolves and deteriorates. It is worth pointing out that the severity of traffic deterioration in terms of traffic delay and road capacity reduction dramatically depends on the traffic blocking duration time.

Thus, the failure duration rate that incorporates failure duration time is proposed for analyzing the LOF of bus stops. The failure duration time can be measured by the timespan (the waiting time of buses outside the stop) for all failed bus stopping services during a given time period. Specifically, it can be calculated by examining the arrival and departure time of buses using the loading areas of a bus stop. Then, the failure duration rate (FDR) can be formulated by

$$R_{fd} = \frac{t_{fd}}{t_s} = \frac{t_{fd}}{t - t_v}, \tag{6}$$

where t_{fd} is the failure duration time (sec), t_s represents the total occupancy time of using the bus stop (sec), and t_v is the duration time for a vacant bus stop (sec).

The FDR can be interpreted as the ratio of waiting and blocking for arriving buses during a given time period. For a specific bus stop failure, the average duration time per failure (t_{fdr}) can be formulated as

$$t_{fdr} = \frac{t_{fd}}{Q_s - \left(Q_s^{n=1} + Q_s^{n=2} + \cdots + Q_s^{n=b}\right)}. \tag{7}$$

Then, the average duration rate per failure (R_{fdr}) is obtained as

$$R_{fdr} = \frac{t_{fdr}}{t - t_v}. \tag{8}$$

3. Characteristic Analysis

3.1. Data Collection.
In this study, bus dwelling data are based on the AVL data provided by Wujiang Transit Agency in Suzhou, China. The AVL data span 56 consecutive working days from October 22, 2018, to January 9, 2019. The dataset of each day has a half-day bus dispatching time window, from 7:00 to 19:00, and there are nine bus routes and eight bay-type bus stops (see their geographic locations in Figure 1). These eight test bus stops keep considerable distances to intersections (the average distance of 200 m), and thus the interaction between the bus bay and nearby intersections would be negligible.

The details of the nine bus routes associated with each bus stop are given in Table 1, and the headways of these transit routes range from 8 to 15 minutes.

The dwelling time of buses serving these 9 routes at the 8 bus stops are extracted from the collected AVL data. Some records are provided in Table 2. Based on the arrival and departure time of buses boarding and alighting passengers at the third bus stop, we can determine the bus failure characteristics. For example, the bus with the ID of SU-EU9353 serving Route 710 departed from the bus stop at 16:14:28, while the bus with the ID of SU-EU6029 serving Route 741 arrived at the same stop at 16:15:07, and the bus arriving later

needed to wait outside the stop for 13 seconds (failure duration time) until bus no.710 left the stop.

3.2. Temporal Characteristics Analysis.
The number and duration of buses stop failure (per hour) at bus stops are determined based on AVL data. We here use equations (5) and (6) to calculate the hourly FR and FDR of the eight bus stops. Figure 2 plots the hourly time-varying characteristics of R_f and R_{fd} over 56 working days of the No. 1 bus stop (672 hourly R_f data and R_{fd} data in total). It can be observed that R_f and R_{fd} in the morning (7:00–9:00) and evening peak hours (17:00–19:00) are higher than those in nonpeak hours. Overall, the mean value of R_f is more significant than that of R_{fd} in most times.

We then look at the median value of the hourly R_f and R_{fd}. As we can see in Figure 3, there are highly similar patterns of R_f and R_{fd} during peak hours for all the eight bus stops. The reason might be that there is high travel demand in both passengers and vehicles are in at peak hours, and thus bus stop failures occur more frequently, especially because of a long time for passengers to board and alight.

3.3. Failure Duration Analysis.
Figure 4 shows the relationship between R_f and R_{fd}. For each bus stop, 224 data during the morning and evening peak hours (7:00 to 9:00 and 17:00 to 19:00) are considered.

From Figure 4, the hourly FDR shows a weak positive correlation to the FR for all the eight bus stops. Overall, the fitting parameters between R_f and R_{fd} vary remarkably among bus stops sites. For example, there are the lowest R-square (0.31) for no.1 bus Stop and the highest R-square (0.661) for no. 4 bus stop. In addition, the statistical relationship between R_f and R_{fd} of 1-berth stops (no.1 through no.6 bus stops) is weaker than that of the two 2-berth stops (no.7 and no.8 bus stops).

As mentioned above, it is difficult to determine a well-fitted failure duration rate function (for the 8 bus stops) using failure rate directly due to the dispersion. Taking no.1 bus stop as an example, R_{fd} is increased from 13.3% to 33.0% as R_f changes from 24% to 25%. But R_f increases from 16.7% to 43.3% as R_{fd} increases from 24% to 25%. It indicates the significant dispersion. To decrease the dispersion of failure duration rate in analysis, the failure rate of bus stops is divided into sections with an interval length of 5%. $R_f^{5\%}$ is defined as the 5% section level of failure rate at a bus stop, which range from 5% to 10% ($R_f^{5\%} \in [5\%, 10\%)$). Then, the mean, standard deviation (S.D.), and coefficient of variation (C.V.) of failure duration rate for these 8 bus stops are clustered and calculated at a different level of failure rate. The mean, S.D., and C.V. of failure duration rates at different failure rate level of these eight test bus stops are compared in Table 3. When the frequency of bus stop failure is less than 5 at a level, the failure duration rate is not calculated and identified as "not available (N/A)" in our analysis. Because the span of these 8 bus stops is different, each bus stop has some "not available (N/A)" at a corresponding failure rate level.
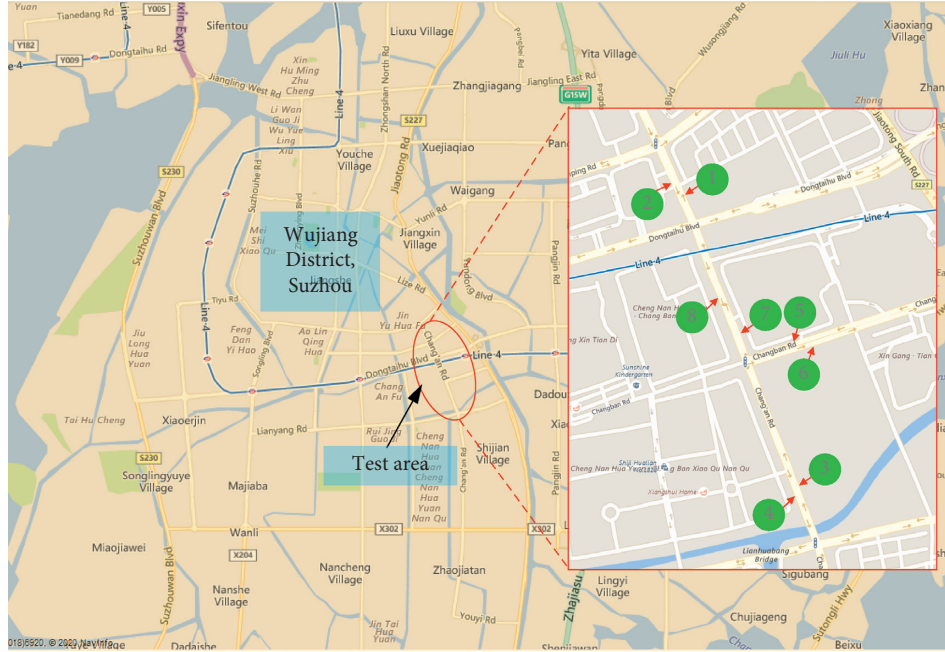
Figure 1: Locations of eight bus stops for AVL data collection.

Table 1: The details of test bus stops.

No.	Bus stop (number of bus berths)	Direction	Bus routes
1	Hengda Market West (1)	S $\longrightarrow$ N	Routes 708, 719, 732, 733
2	Hengda Market West (1)	N $\longrightarrow$ S	Routes 708, 719, 732, 733
3	Changan Rd. and Lianyang Rd. North (1)	S $\longrightarrow$ N	Routes 701, 710, 711, 720, 741
4	Changan Rd. and Lianyang Rd. North (1)	N $\longrightarrow$ S	Routes 701, 710, 711, 720, 741
5	Changban Rd. and Changan Rd. East (1)	E $\longrightarrow$ W	Routes 701, 710, 711
6	Changban Rd. and Changan Rd. East (1)	W $\longrightarrow$ E	Routes 701, 710, 711
7	Changan Rd. and Changban Rd. North (2)	S $\longrightarrow$ N	Routes 708, 719, 720, 732, 733, 741
8	Changan Rd. and Changban Rd. North (2)	N $\longrightarrow$ S	Routes 708, 719, 720, 732, 733, 741

Table 2: Some buses arrival and departure time records extracted from AVL data.

Route no.	Bus ID	Bus stop	Date/time	
			Arrival time	Departure time
701	SU-EU9526	No.3 bus stop	2018-12-18 16:10:31	2018-12-18 16:11:11
710	SU-EU9353	No.3 bus stop	2018-12-18 16:13:34	2018-12-18 16:14:28
741	SU-EU6029	No.3 bus stop	2018-12-18 16:14:15	2018-12-18 16:15:07
720	SU-EU9582	No.3 bus stop	2018-12-18 16:19:15	2018-12-18 16:20:08
701	SU-EU9379	No.3 bus stop	2018-12-18 16:20:30	2018-12-18 16:21:02
711	SU-EU9359	No.3 bus stop	2018-12-18 16:22:16	2018-12-18 16:23:19

Based on the mean, S.D., and C.V. of failure duration rates at different failure rate level, we then analyze the failure duration characteristics. Figure 5 presents the average value (for all these eight bus stops) of mean, S.D., and C.V. for failure duration rate at different failure rate levels. The mean and S.D. of R_{fd} increase with the increasing failure rate level, and the mean has a higher and faster growth rate than that of S.D. The mean of R_{fd} is increased from 5.2% to 27.5% as failure rate level increases from 5% to 40%. And the S.D. of R_{fd} is increased from 3.1% to 5.5% as the same span of failure rate level.

However, the favorable trend does not hold for the C.V. of R_{fd}, as illustrated in Figure 5. The C.V. of R_{fd} is decreased from 59.6% to 19.9% as failure rate level changes from 5% to 40%.

Figure 6 displays the average value of the mean, S.D., and C.V. for failure duration rate of six 1-berth stops and two 2-berth stops. The average value of the mean, S.D., and C.V. for failure duration rate of the single-berth stops is more significant than that of the 2-berth stops.

Figure 7 displays the correlation relationships between FDR per failure (R_{fdr}) and the FR for the eight bus stops. It

(a)

(b)
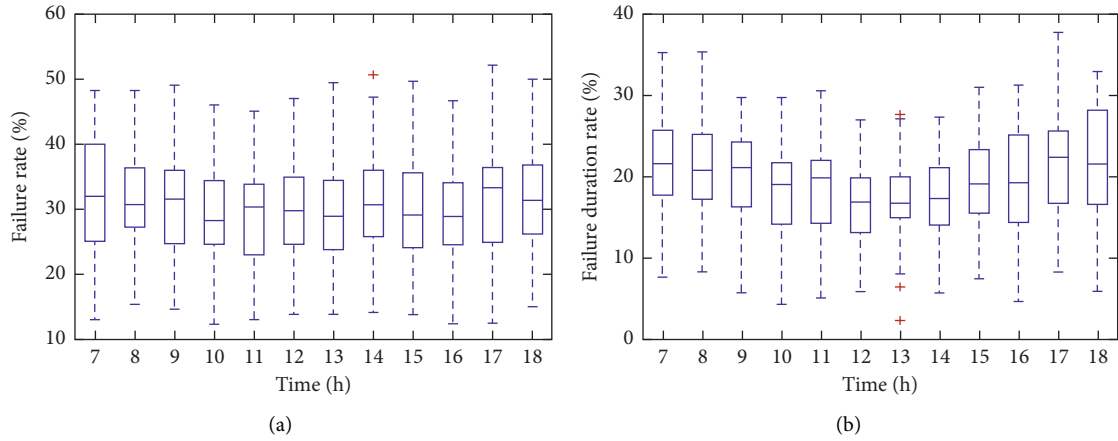
FIGURE 2: The time-varying characteristics of no.1 bus stop failure.

(a)
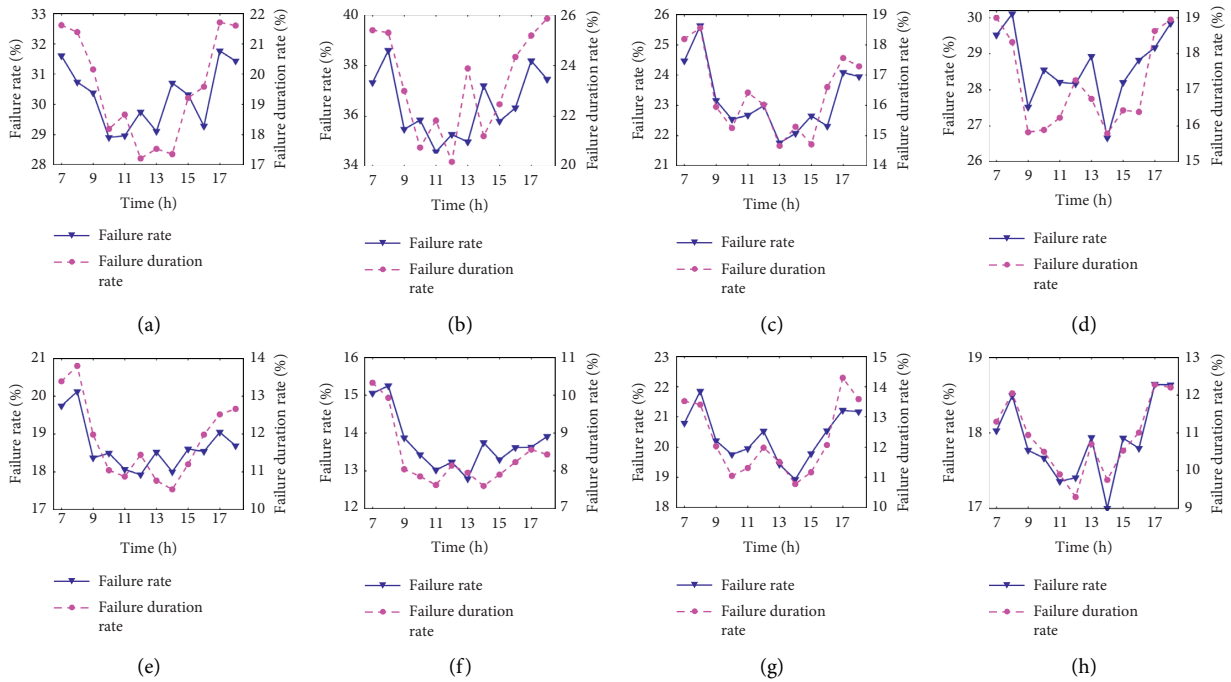
(b)

(c)

(d)

(e)

(f)

(g)

(h)

FIGURE 3: Bus stop failure time-varying characteristics of test bus stops. (a) No.1 bus stop. (b) No.2 bus stop. (c) No.3 bus stop. (d) No.4 bus stop. (e) No.5 bus stop. (f) No.6 bus stop. (g) No.7 bus stop. (h) No.8 bus stop.
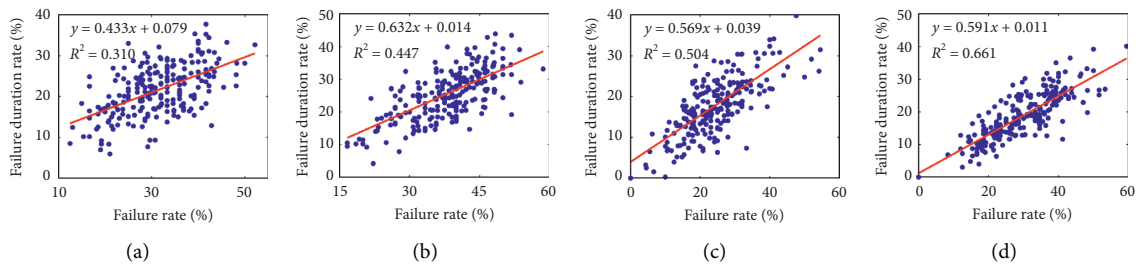

(a)
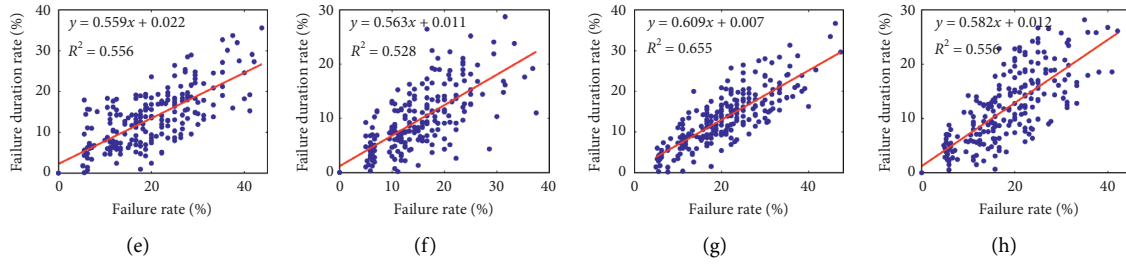
(b)

(c)

(d)

FIGURE 4: Continued.

FIGURE 4: The relationship analysis between FDR vs. FR. (a) No.1 bus stop. (b) No.2 bus stop. (c) No.3 bus stop. (d) No.4 bus stop. (e) No.5 bus stop. (f) No.6 bus stop. (g) No.7 bus stop. (h) No.8 bus stop.

TABLE 3: Failure duration rate of test bus stops at different failure rate levels.

| Failure rate level		No.1 bus stop	No.2 bus stop	No.3 bus stop	No.4 bus stop	No.5 bus stop	No.6 bus stop	No.7 bus stop	No.8 bus top
$R_f^{5\%}$	Mean	N/A	N/A	N/A	N/A	0.0710	0.0504	0.0382	0.0464
	S.D.	N/A	N/A	N/A	N/A	0.0508	0.0289	0.0225	0.0236
	C.V.	N/A	N/A	N/A	N/A	0.7155	0.5732	0.5878	0.5081
$R_f^{10\%}$	Mean	N/A	N/A	0.0896	0.0885	0.0960	0.0807	0.0835	0.0797
	S.D.	N/A	N/A	0.0375	0.0360	0.0431	0.0383	0.0397	0.0353
	C.V.	N/A	N/A	0.4188	0.4070	0.4488	0.4750	0.4750	0.4430
$R_f^{15\%}$	Mean	0.1443	N/A	0.1541	0.1104	0.1155	0.1133	0.1118	0.1070
	S.D.	0.0502	N/A	0.0544	0.0355	0.0480	0.0520	0.0310	0.0461
	C.V.	0.3477	N/A	0.3527	0.3218	0.4159	0.4587	0.2769	0.4306
$R_f^{20\%}$	Mean	0.1735	0.1380	0.1630	0.1405	0.1338	0.1304	0.1422	0.1450
	S.D.	0.0498	0.0554	0.0510	0.0378	0.0494	0.0499	0.0429	0.0510
	C.V.	0.2871	0.4015	0.3129	0.2688	0.3688	0.3828	0.3015	0.3516
$R_f^{25\%}$	Mean	0.2104	0.2075	0.1904	0.1860	0.1776	0.1664	0.1718	0.1754
	S.D.	0.0540	0.0661	0.0532	0.0450	0.0533	0.0576	0.0430	0.0497
	C.V.	0.2566	0.3186	0.2794	0.2417	0.3002	0.3459	0.2505	0.2833
$R_f^{30\%}$	Mean	0.2231	0.2159	0.2236	0.1975	0.1698	0.1832	0.1939	0.1780
	S.D.	0.0495	0.0558	0.0554	0.0370	0.0402	0.0610	0.0460	0.0470
	C.V.	0.2217	0.2583	0.2479	0.1872	0.2368	0.3328	0.2373	0.2640
$R_f^{35\%}$	Mean	0.2215	0.2506	0.2569	0.2421	0.2641	N/A	0.2282	0.2328
	S.D.	0.0478	0.0539	0.0546	0.0458	0.0608	N/A	0.0405	0.0470
	C.V.	0.2160	0.2151	0.2125	0.1890	0.2304	N/A	0.1773	0.2019
$R_f^{40\%}$	Mean	0.2719	0.2828	0.2885	0.2587	N/A	N/A	N/A	N/A
	S.D.	0.0532	0.0609	0.0512	0.0543	N/A	N/A	N/A	N/A
	C.V.	0.1957	0.2154	0.1774	0.2098	N/A	N/A	N/A	N/A
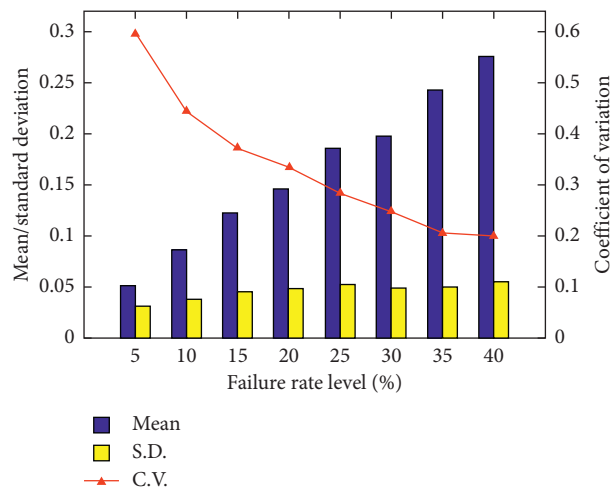
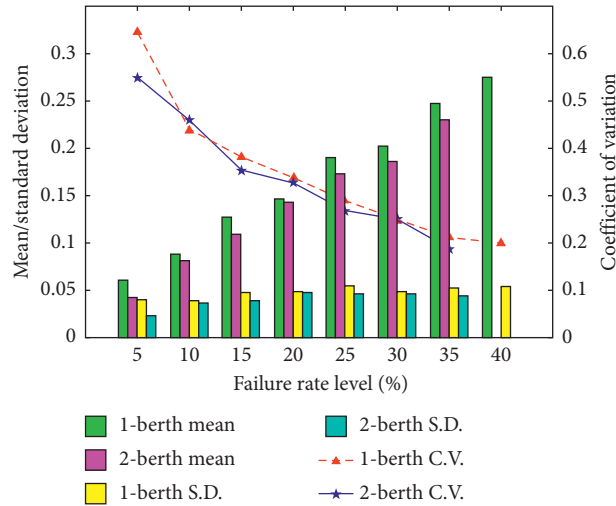FIGURE 5: Average mean, S.D., and C.V. of FDR for eight bus stops.

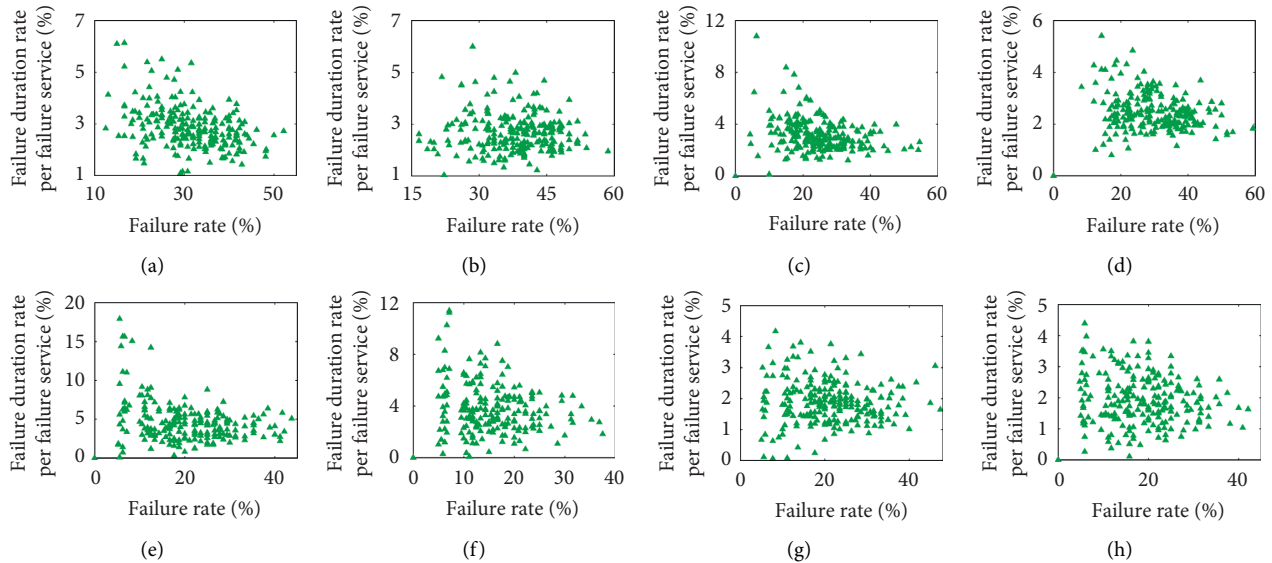Figure 6: Average mean, S.D., and C.V. of FDR for two types of bus stops.

Figure 7: FDR per failure vs. FR for test bus stops. (a) No.1 bus stop. (b) No.2 bus stop. (c) No.3 bus stop. (d) No.4 bus stop. (e) No.5 bus stop. (f) No.6 bus stop. (g) No.7 bus stop. (h) No.8 bus stop.

is shown that the FR is insensitive to R_{fdr}. For the no.1 bus stop, Figure 7(a) reveals that R_{fdr} is increased from 1.5% to 5.4% and from 1.9% to 3.4% when the FR climbs up from 21% to 22% and from 34% to 35%, respectively. Thus, the span of R_{fdr} is significant at the low level of failure rate, and the volatility of R_{fdr} becomes more and more slight with the increasing failure rate. Besides, the diversity correlation relationship between R_{fdr} and R_f is influenced by the number of berths of loading areas. Figure 7 shows that R_{fdr} is increased from 1.2% to 3.8% and from 1.5% to 5.4%, respectively, as the failure rate increase from 21% to 22% for no.7 bus stop (2-berth type) and no.1 bus stop (1-berth type). Therefore, compared with single-berth bus stops, the two 2-berth bus stops have less R_{fdr} at the same level of failure rate.

For the different levels of failure rate (range from 5% to 40%), the value of the mean, S.D., and C.V. for R_{fdr} at the

bus stops are calculated and presented in Figure 8. The average values of the mean, S.D., and C.V. for R_{fdr} reveal a significant negative correlation with the FR level. When the FR ranges from 30% to 40%, the three statistics for R_{fdr} are not very sensitive to the FR level, and there are no obvious fluctuations. For the mean value of R_{fdr}, the maximum variation is merely 0.1% when the FR level falls into the range of 30% to 40%. Therefore, the dispersion of R_{fdr} decreases sharply with the FR level (especially when the FR is greater than 30%), which implies that the average FDR per failed is comparatively stable. In Figure 9, there is a similar trend of C.V.

For the different levels of failure rate (range from 5% to 40%), the value of the mean, S.D., and C.V. for R_{fdr} for six 1-berth stops and two 2-berth stops are calculated and depicted in Figure 9. The three statistics for R_{fdr} at different
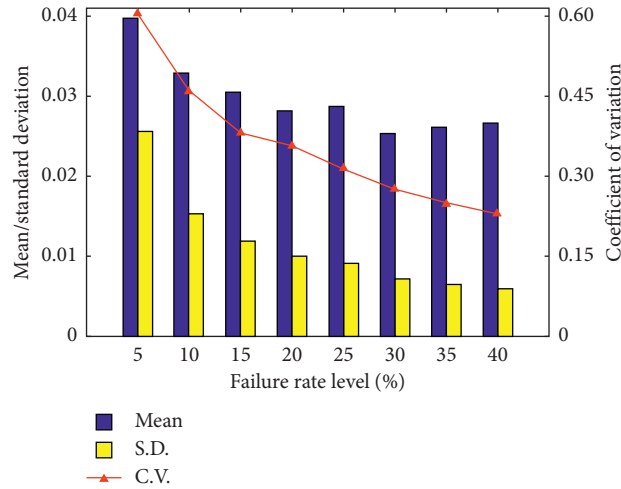
Figure 8: Average mean, S.D., and C.V. of FDR per failure for eight bus stops.
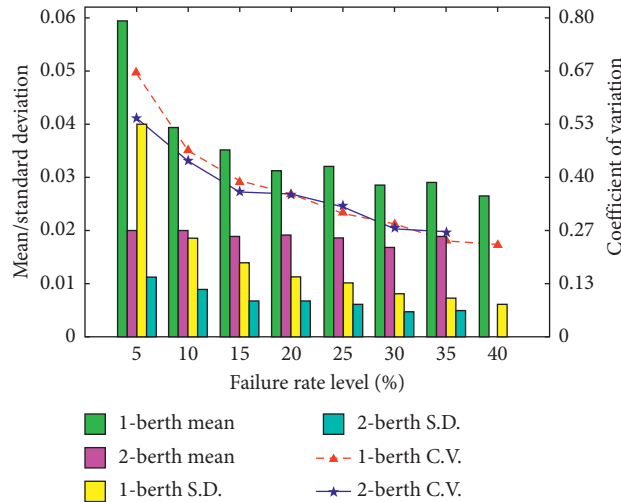

Figure 9: Average mean, S.D., and C.V. of FDR per failure for two types of bus stops.

failure rate levels for the single-berth stops is larger than that of the 2-berth stops, which shows higher stability at the 2-berth bus stops.

4. Correlation Analysis

As discussed in Section 3, it is found that it is not so easy to establish a satisfactory relationship between the FDR and FR, $f(R_{fd}, R_f)$, via linear regression models. In this section, we propose a "Distribution Fitting and Cumulative Distribution Correlation (DF-CDC)" method for an in-depth and more reasonable analysis of the correlation between R_{fd} and R_f.

4.1. Distribution Fitting. The distribution fitting analysis is regarded as a useful approach to mine characteristics of transit operational parameters from the probabilistic perspective [33]. A unified probability distribution is explored that can well fit the essence of R_f and R_{fd}, utilizing the

probability and statistical analysis methods. That is, we aim to understand how well a candidate distribution is fitted with predicted parameters for R_f and R_{fd}. Typically, chi-squared (χ^2), Kolmogorov–Smirnov (K-S), and Anderson–Darling (A-D) tests could be used for assessing the goodness-of-fit of our analysis results. In this paper, the K-S test statistic at a significance level of 0.05 is adopted for the goodness-of-fit test based on the data of 8 bus stops on peak periods per workday (224 data per stop). 36 probability distributions listed in Table 4 are chosen for hypothesis analysis. Table 4 shows the number of rejections for the 36 possible candidates. The results reveal that the distributions of Error, Gen. Extreme Value, Gen. Logistic, Logistic, and Normal could be selected as the candidate distributions for the correlation analysis between R_{fd} and R_f.

The five well-fitted candidate probability distributions (Error distribution, Gen. Extreme Value distribution, Gen. Logistic distribution, Logistic distribution, and Normal distribution) are redeemed for analyzing the goodness-of-fit of fitted FR and FDR at the bus stops. After estimating the

TABLE 4: Candidate distributions fitting results for bus failure characteristics.

No.	Distribution	Number of rejections	
		FR	FDR
1	Beta	4	4
2	Burr	2	1
3	Cauchy	7	6
4	Dagum	2	0
5	Erlang	5	3
6	Error	0	0
7	Fatigue Life	3	4
8	Gamma	2	0
9	Gen. Extreme Value	0	0
10	Gen. Gamma	1	0
11	Gen. Logistic	0	0
12	Gen. Pareto	2	1
13	Gumbel max	6	2
14	Gumbel min	7	6
15	Hypersecant	2	2
16	Inv. Gaussian	4	3
17	Johnson SB	2	0
18	Johnson SU	6	8
19	Kumaraswamy	4	4
20	Laplace	7	6
21	Logistic	0	0
22	Log-Logistic	3	1
23	Lognormal	3	4
24	Log-Pearson 3	5	6
25	Nakagami	1	0
26	Normal	0	0
27	Pearson 5	5	6
28	Pearson 6	2	1
29	Pert	4	2
30	Phased Biexponential	8	6
31	Phased Bi-Weibull	3	1
32	Rayleigh	4	5
33	Rice	2	1
34	Triangular	4	4
35	Uniform	4	2
36	Weibull	2	1

Note. 5 distributions highlighted in grey shading are selected as the well-fitted candidate distribution of R_f and R_{fd}.

parameters of these distributions (using the Probability density function for distributions shown in Table 5), K-S test results (P value) for the five candidate probability distributions are plotted in Figure 10. It can be seen that the goodness-of-fit measured by P values for FDR fitted distribution is much better than that for the FR.

In Table 5, the means of P value for R_f and R_{fd} at the eight bus stops for the five candidate distributions are also given. Gen. Extreme Value distribution is the best one in terms of P value (with 0.68224 of FR and 0.87865 of FDR) in distribution fitting for the R_f and R_{fd}.

4.2. Cumulative Distribution Correlation.
The probability density function of Gen. Extreme Value distribution is utilized for fitting the hourly FR and FDR distributions for the eight bus stops. The results are provided in Table 6.

Based on the calculated parameters of the fitted Gen. Extreme Value distribution in Table 6, the cumulative

distribution function (CDF) curve of R_f and R_{fd} for these test bus stops can be determined. From the fitted CDF curve, the fitted value at a different level of CDF can be recorded. The fitted value at a different level of CDF can be recorded using the fitted CDF curve, and the actual value can be determined by analyzing the ranking level based on the sorted 224 data collected at each test bus stop. For analyzing the accuracy of the fitted CDF value, the relative error between the actual and fitted value of R_f and R_{fd} for test bus stops at a different level of CDF is examined. Furthermore, 17 critical levels of CDF, ranging from 10% to 90% (with 5% of interval length), are selected for verifying. The relative error between the actual and fitted value of R_f and R_{fd} for the critical CDF level at the 8 test bus stops are presented in Tables 7 and 8.

Expect for a tiny minority of critical CDF level, relative errors between the actual and fitted value of R_f and R_{fd} are always less than 10% at these test bus stops, as shown in Tables 7 and 8. Therefore, Gen. Extreme Value distribution

TABLE 5: Average P value of FR and FDR at test bus stops for well-fitted candidate distributions.

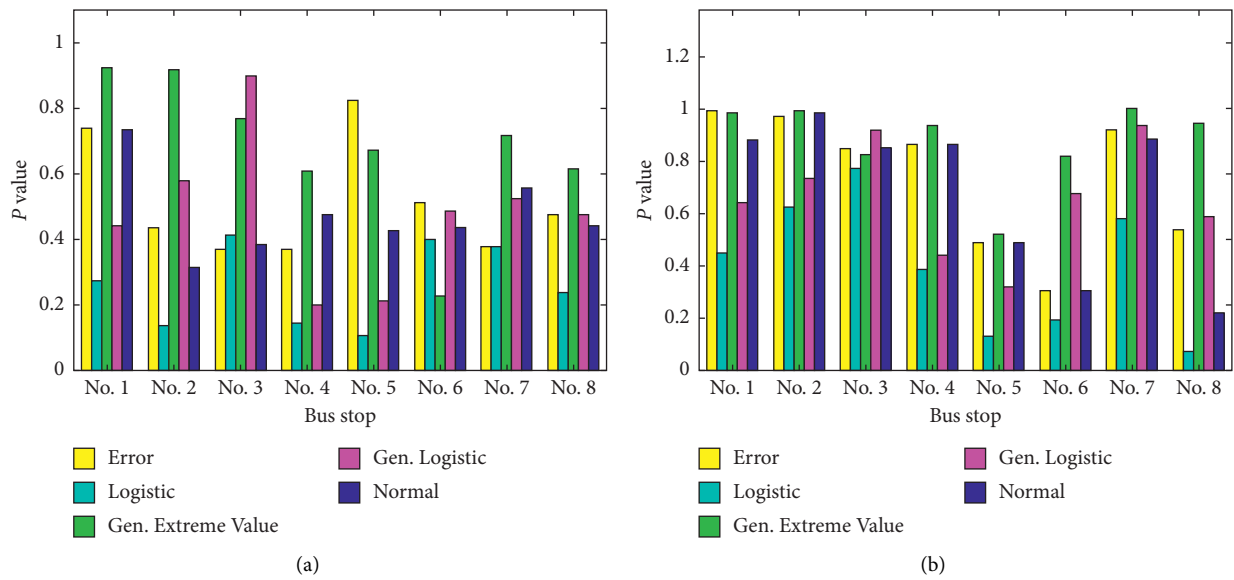| Distribution	Probability density function (PDF) and sample space	Average P value	
		FR	FDR
Error	$f(x \mid \sigma, k, \xi) = c_1 \sigma^{-1} \exp(-\|c_0 z\|^k) c_0 = (\Gamma(3/k)/\Gamma(1/k))^{1/2} \ c_1 = kc_0/2\Gamma(1/k) \ z \equiv (x-\xi)/\sigma \sigma > 0,$ $\xi \le x < +\infty$	0.51376	0.74114
Gen. Extreme Value	$f(x \mid \sigma, k, \xi) = \begin{cases} \exp(-(1+kz)^{-1/k}(1+kz)^{(k+1/-k)})/\sigma, & k \neq 0, \\ \exp(-z - \exp(-z))/\sigma, & k = 0, \end{cases} \ \sigma > 0, \ \xi \le x < +\infty$	0.68224	0.87865
Gen. Logistic	$f(x \mid \sigma, k, \xi) = \begin{cases} (1+kz)^{-1-1/k}/\sigma(1+(1+kz)^{-1/k})^2, & k \neq 0, \\ \exp(-z)/\sigma(1+\exp(-z))^2, & k = 0, \end{cases} \ \sigma > 0, \ \xi \le x < +\infty$	0.47726	0.65518
Logistic	$f(x \mid \sigma, \xi) = \exp(-z)/\sigma(1+\exp(-z))^2 \sigma > 0, \ \xi \le x < +\infty$	0.25996	0.40434
Normal	$f(x \mid \sigma, \xi) = \exp(-(1/2)(x - \xi/\sigma)^2)/\sigma\sqrt{2\pi}\sigma > 0, \ \xi \le x < +\infty$	0.47032	0.68553

FIGURE 10: Well-fitted candidate distributions P value comparison for bus failed characteristics. (a) Failure rate (%). (b) Failure duration rate (%).

TABLE 6: Parameters of fitted FR and FDR at test bus stops.

Bus stop	Fitted FR parameters			Fitted FDR parameters		
	σ	k	ξ	σ	k	ξ
No.1 bus stop	−0.3034	0.08116	0.28627	−0.30432	0.06281	0.19447
No.2 bus stop	−0.42799	0.08592	0.35594	−0.31185	0.07768	0.22823
No.3 bus stop	−0.14816	0.08291	0.20818	−0.21402	0.07047	0.15087
No.4 bus stop	−0.28909	0.10029	0.26163	−0.25443	0.07185	0.16056
No.5 bus stop	−0.24515	0.09676	0.15746	−0.18459	0.06912	0.10188
No.6 bus stop	−0.22178	0.07511	0.11501	−0.11238	0.05439	0.06722
No.7 bus stop	−0.20209	0.08564	0.17765	−0.1715	0.06254	0.1101
No.8 bus stop	−0.21955	0.08123	0.15239	−0.13328	0.06037	0.09181

performs well in fitting CDF value for R_f and R_{fd} at these test bus stops, and the accuracy and reliability for the fitted value are convincing.

For the distribution pattern of R_f and R_{fd} at test bus stops, Gen. Extreme Value distribution can be well fitted. Moreover, they also have a positive correlation relationship (as shown in Figure 4). Therefore, it can be considered that the fitted value of R_f and R_{fd} at the same CDF level perform equivalently in

failure characteristics analysis. Figure 11 presents the fitted curves and critical level values of R_f and R_{fd}, which are fitted and utilized Gen. Extreme Value distribution for no.1 bus stop. In Figure 11, the "star (with pink color)" and "triangle (with red color)" display the fitted values of R_f and R_{fd} at these 17 critical CDF levels, and the "star" and "triangle" connected by a dotted line is defined as an equal correlation pair of R_f and R_{fd} for the corresponding critical CDF level.

TABLE 7: Relative error between actual and fitted FR for test bus stops at different CDF levels.

Level of CDF	No.1 bus stop	No.2 bus stop	No.3 bus stop	No.4 bus stop	No.5 bus stop	No.6 bus stop	No.7 bus stop	No.8 bus stop
$P_c^{10\%}$	0.72	0.22	4.05	3.62	8.67	10.43	0.15	32.67
$P_c^{15\%}$	3.01	2.43	1.91	3.18	26.52	1.33	7.19	1.37
$P_c^{20\%}$	1.97	0.08	3.13	1.02	2.24	15.97	7.90	0.48
$P_c^{25\%}$	0.95	1.37	0.76	2.27	5.87	10.43	0.84	0.09
$P_c^{30\%}$	1.85	1.80	0.33	3.93	5.63	4.26	3.13	4.10
$P_c^{35\%}$	0.55	1.26	2.02	0.46	0.28	5.36	1.70	3.53
$P_c^{40\%}$	0.55	0.09	0.07	0.39	0.50	2.79	1.30	0.97
$P_c^{45\%}$	2.32	0.17	0.33	0.83	1.23	1.38	1.76	2.13
$P_c^{50\%}$	0.09	0.43	1.22	1.66	2.07	0.99	1.25	0.45
$P_c^{55\%}$	1.18	1.10	0.30	0.25	2.11	1.01	1.23	4.05
$P_c^{60\%}$	0.68	1.52	0.75	3.14	2.18	2.87	0.29	1.56
$P_c^{65\%}$	0.12	0.42	0.74	2.39	1.79	3.62	3.20	3.32
$P_c^{70\%}$	0.52	1.06	2.66	1.71	1.76	3.77	2.94	2.30
$P_c^{75\%}$	0.03	0.25	0.83	0.03	2.00	1.61	2.09	1.19
$P_c^{80\%}$	1.85	0.57	0.47	0.35	0.40	1.21	1.92	2.48
$P_c^{85\%}$	0.09	0.63	3.55	2.68	1.77	2.30	0.06	2.79
$P_c^{90\%}$	0.47	0.38	0.55	3.63	3.94	4.19	0.25	1.69

Note: $P_c^{10\%}$ represents the 10% CDF level (the 10th percentile of the fitted FR), ranking in ascending order ranging from 0 to 1.

TABLE 8: Relative error between actual and fitted FDR for test bus stops at different CDF levels.

Level of CDF	No.1 bus stop (%)	No.2 bus stop (%)	No.3 bus stop (%)	No.4 bus stop (%)	No.5 bus stop (%)	No.6 bus stop (%)	No.7 bus stop (%)	No.8 bus stop (%)
$P_c^{10\%}$	1.60	0.46	5.16	1.94	12.48	34.54	4.82	0.89
$P_c^{15\%}$	0.68	0.88	0.32	2.49	4.14	11.60	3.99	5.04
$P_c^{20\%}$	0.83	2.61	2.38	1.82	7.78	0.14	0.64	8.65
$P_c^{25\%}$	1.84	2.32	3.42	2.79	7.60	0.85	0.64	3.29
$P_c^{30\%}$	2.31	1.04	0.36	1.29	9.65	3.32	0.08	1.98
$P_c^{35\%}$	0.90	0.09	1.12	1.11	3.08	1.17	2.03	1.43
$P_c^{40\%}$	0.34	1.09	0.17	0.29	0.40	1.06	1.84	3.81
$P_c^{45\%}$	0.77	0.23	0.76	0.27	2.88	3.61	1.02	5.75
$P_c^{50\%}$	0.48	0.62	0.42	2.56	4.19	2.81	1.47	2.73
$P_c^{55\%}$	0.48	1.36	0.76	1.97	4.28	2.28	0.36	1.36
$P_c^{60\%}$	1.32	0.29	1.06	1.08	2.90	1.33	1.05	0.21
$P_c^{65\%}$	0.96	0.69	0.48	1.90	0.52	2.10	1.59	1.61
$P_c^{70\%}$	0.44	1.32	0.17	2.14	1.11	0.47	1.76	0.04
$P_c^{75\%}$	0.32	0.85	1.21	2.65	0.16	0.11	1.78	0.45
$P_c^{80\%}$	0.02	0.27	2.47	0.51	3.55	0.04	0.15	3.59
$P_c^{85\%}$	0.05	0.75	3.98	2.37	7.96	3.19	1.56	0.56
$P_c^{90\%}$	1.51	0.42	0.31	3.62	4.07	0.65	1.93	1.04

Figure 12 reveals the relationship for the 136 equal correlation pair of R_f and R_{fd}, including the pairs for 17 critical CDF levels at 8 test bus stops. A quite strong linear regression expression (with 0.98 of R-square) for the fitted couples of R_f and R_{fd} can be observed, which can reflect the significant correlation relationship between fitted R_f and R_{fd}.

4.3. Correlation Performance Evaluation. The cross-validation method [34] for R_{fd} prediction based on AVL data is adopted for analyzing the correlation between R_{fd} and R_f. A four-step procedure for predicting R_{fd} at a certain cumulative distribution ranking level is illustrated as below. Firstly, the observed R_{fd} and R_f in peak hours of 56 workdays at 75% bus stops of the total 8 test bus stops are selected randomly as the modeling datasets, and the data of the rest two bus stops are defined as predicted datasets. Secondly, the probability density distribution of R_{fd} and R_f (for the selected six bus stops in modeling datasets) are fitted using Gen. Extreme value distribution, and the fitted R_{fd} and R_f are recorded based on their probability density functions at critical cumulative distribution levels, respectively. Thirdly, linear regression expression is developed based on these fitted R_{fd} and R_f for the six bus stops (in modeling dataset) at critical cumulative distribution levels. Finally, the observed R_f (for two bus stops in prediction dataset) at corresponding critical cumulative distribution levels are determined, and the predicted R_{fd} can be calculated using the linear regression model (as formulated in Step 3). The prediction accuracy and reliability of R_{fd} can be determined by comparing the actual and predicted value.
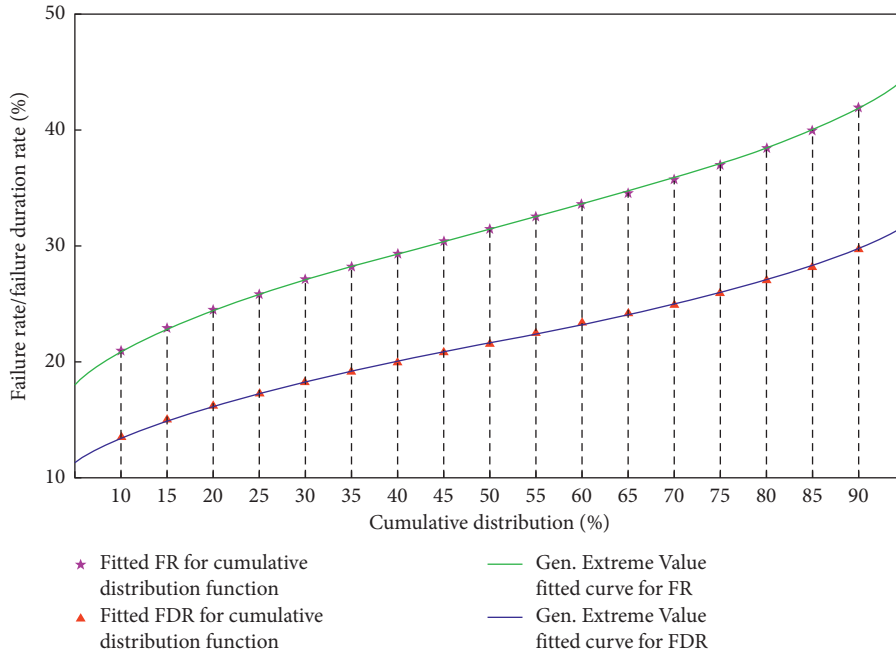
FIGURE 11: Correlation between fitted FR and FDR at the same CDF level using Gen. Extreme Value distribution.
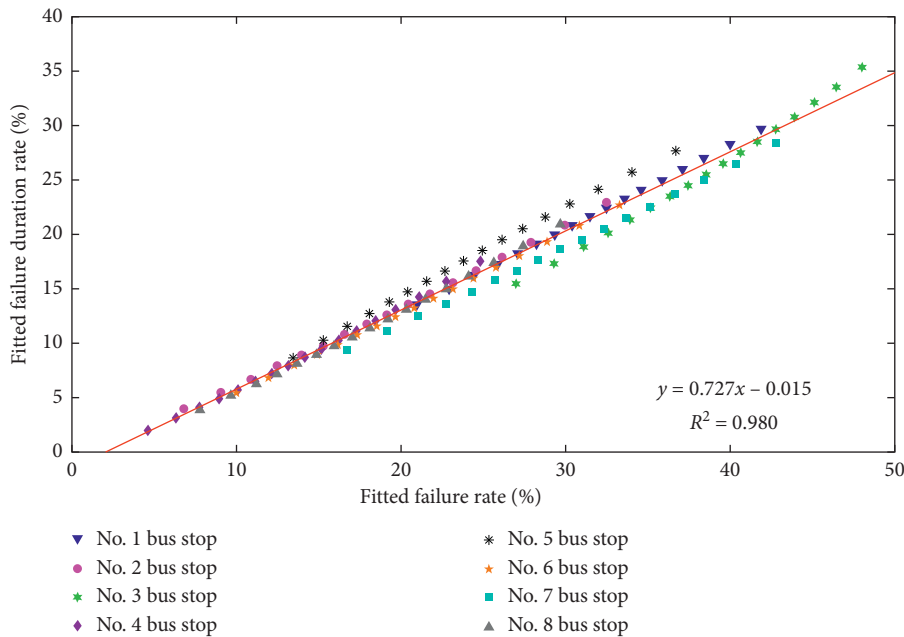

FIGURE 12: Gen. Extreme Value distribution fitted FDR vs. FR correlation analysis.

It is obvious that there are 28 different combinations of modeling and prediction datasets, as the rules described in Step 2. There are seven estimated values of R_{fd} for different cumulative distribution levels at each bus stop, calculated from these 28 different division plans. Based on cross-validation, the predicted value and relative error of R_{fd} for different cumulative distribution levels at 8 test bus stops are calculated, and the results are shown in Figure 13.

In Figure 13, the mean value of predicted relative errors for seven times R_{fd} prediction at different cumulative distribution levels is represented by a solid blue line. According to the results, these test bus stops have more accurate predicted values (the relative error is less than 15%) at most cumulative distribution levels, expect for low cumulative distribution levels (less than 15%). The prediction results of R_{fd} for different cumulative distribution levels at no.1 bus stop perform well (the relative error is not more than 5%) in general. Also, for most of the bus stops (no.2, no.4, no.5, no.6, no.7, and no.8 bus stop), the relative error of predicted R_{fd} is diminished gradually, as the cumulative
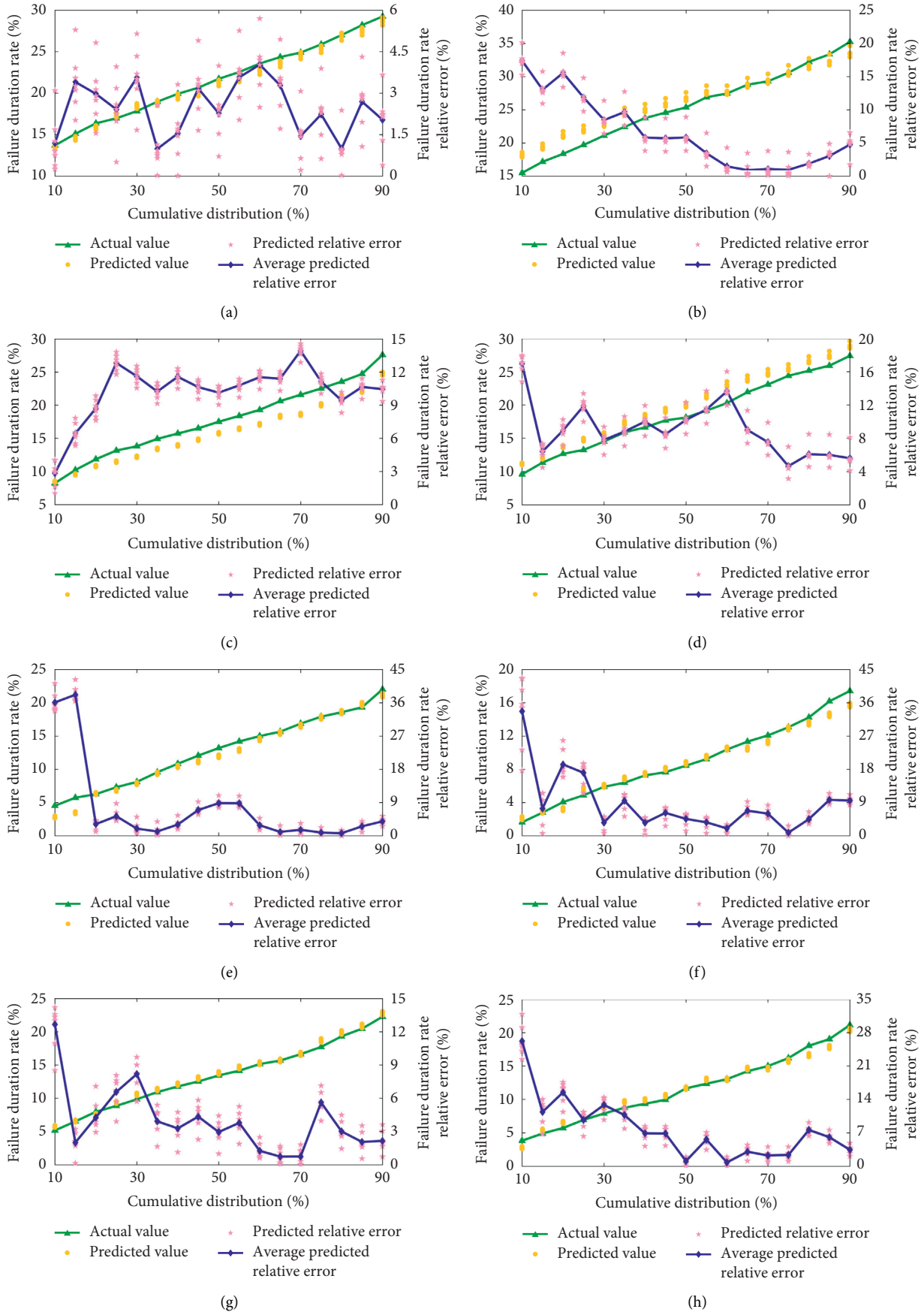
Figure 13: Results of FDR estimation for test bus stops at CDF levels. (a) No.1 bus stop. (b) No.2 bus stop. (c) No.3 bus stop. (d) No.4 bus stop. (e) No.5 bus stop. (f) No.6 bus stop. (g) No.7 bus stop. (h) No. 8 bus stop.
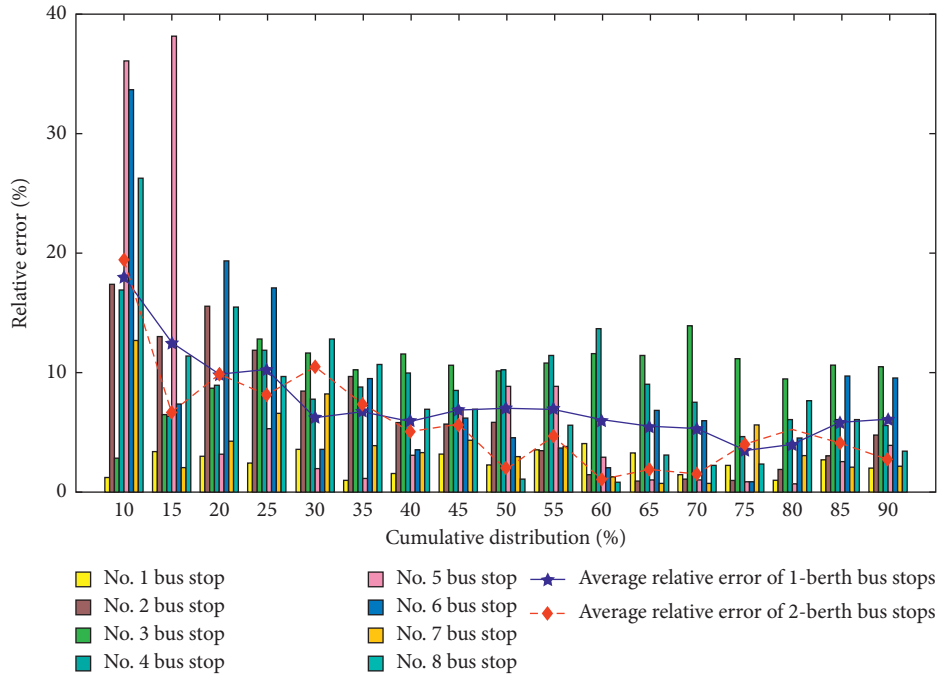
FIGURE 14: Comparison of FDR estimation relative error at CDF levels.

distribution sort decreases. No. 3 and no. 4 bus stops serve five bus routes, more than other bus stops. In general, the more bus routes are served at a bus stop, the more complicated buses arrival patterns are. Therefore, the average predicted relative error values of no. 3 and no. 4 bus stops will be larger than other bus stops.

To make an easy comparison, the average relative error of R_{fd} for 1-berth bus stops and 2-berth bus stops are plotted in Figure 14. It can be observed that, for both types of stops, the average relative error shows a decreasing trend. In addition, the volatility is insensitive when the cumulative distribution level is higher than 30%, and the average relative error is less than 8% in general. According to the results, the proposed "Distribution Fitting and Cumulative Distribution Correlation (DF-CDC)" can develop a significant correlation relationship between the failure rate and failure duration rate.

5. Conclusion

In order to analyze the characteristics of bus stop failure, we propose a new measurement called FDR and make a comparison to the traditional index of the FR. Compared with the FR, the proposed FDR is capable to quantitatively assess the impact of bus stop failure on traffic efficiency. Based on the collected AVL data associated with the eight bus stops in Wujiang District of Suzhou, we make an in-depth analysis of the characteristics and correlation of the FR and FDR. Some insightful findings are summarized as follows:

(i) It can be observed that the values of FR and FDR in morning and evening peak hours are greater than those during off-peak hours. The value of the FR is usually larger than that of the FDR across all the eight bus stops.

(ii) It is found that there is a positive correlation between the FR and the FDR. However, the R-square values of the linear regressions fluctuate dramatically among different bus stops. The results also indicate that the FDR is more robust than the FR in describing transit system status and traffic characteristics.

(iii) We find that Gen. Extreme Value distribution could be well used for the fittings of both the FR and FDR and the proposed "Distribution Fitting and Cumulative Distribution Correlation (DF-CDC)" method works well in determining the fitted values of the FR and FDR at the critical CDF level that reflects a significant correlation between the FR and FDR.

Future works could be extended in two aspects. First, more AVL and other source data in other cities could be collected and used for analyzing the failure characteristics of bus stops. Second, based on the data-driven analysis of bus stop failure characteristics, we could make some scenario analysis to find out the most important factors, such as the number of bus stop berths and the number of lanes or the passenger demand.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

The authors confirm contribution to the paper as follows. R. Li and H. Wang conceptualized and designed the study; X. Xue carried out data collection; R. Li, X. Xue, and H. Wang carried out analysis and evaluation; R. Li and X. Xue prepared draft manuscript. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments

References

[1] KFH Group, *Transit Capacity and Quality of Service Manual*, Transportation Research Board, National Research Council, Washington, DC, USA, 3rd edition, 2013.

[2] J. M. Bunker, "High volume bus stop upstream average waiting time for working capacity and quality of service," *Public Transport*, vol. 10, pp. 311–333, 2018.

[3] A. Tirachini, "Bus dwell time: the effect of different fare collection systems, bus floor level and age of passengers," *Transportmetrica A: Transport Science*, vol. 9, pp. 28–49, 2013.

[4] Q. Meng and X. Qu, "Bus dwell time estimation at bus bays: a probabilistic approach," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 61–71, 2013.

[5] B. Bian, N. Zhu, S. Ling, and S. Ma, "Bus service time estimation model for a curbside bus stop," *Transportation Research Part C: Emerging Technologies*, vol. 57, pp. 103–121, 2015.

[6] H. Yu, Z. Wu, D. Chen, and X. Ma, "Probabilistic prediction of bus headway using relevance vector machine regression," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, pp. 1772–1781, 2017.

[7] M. Zhang, Q. Meng, L. Kang, and W. Li, "Tailored Wakeby-type distribution for random bus headway adherence ratio," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 220–244, 2018.

[8] W. Gu, Y. Li, M. J. Cassidy, and J. B. Griswold, "On the capacity of isolated, curbside bus stops," *Transportation Research Part B: Methodological*, vol. 45, pp. 714–723, 2011.

[9] M. Shen, W. Gu, S. Hu, and H. Cheng, "Capacity approximations for near- and far-side bus stops in dedicated bus lanes," *Transportation Research Part B: Methodological*, vol. 125, pp. 94–120, 2019.

[10] R. Widanapathiranage, J. M. Bunker, and A. Bhaskar, "Modelling the BRT station capacity and queuing for all stopping busway operation," *Public Transport*, vol. 7, pp. 21–38, 2015.

[11] B. Bian, M. Pinedo, N. Zhu, and S. Ma, "Performance analysis of overtaking maneuvers at bus stops with tandem berths," *Transportation Science*, vol. 53, pp. 597–618, 2019.

[12] C. F. Daganzo, "A headway-based approach to eliminate bus bunching: systematic analysis and comparisons," *Transportation Research Part B: Methodological*, vol. 43, pp. 913–921, 2009.

[13] W. Wu, R. Liu, and W. Jin, "Modelling bus bunching and holding control with vehicle overtaking and distributed passenger boarding behavior," *Transportation Research Part B: Methodological*, vol. 104, pp. 175–197, 2017.

[14] C. Wang, Z. Ye, E. Chen, M. Xu, and W. Wang, "Diffusion approximation for exploring the correlation between failure rate and bus stop operation," *Transportmetrica A: Transport Science*, vol. 15, pp. 1306–1320, 2019.

[15] R. Fernández, "Modeling public transport stops by microscopic simulation," *Transportation Research Part C: Emerging Technologies*, vol. 18, pp. 856–868, 2010.

[16] W. Gu, M. J. Cassidy, and Y. Li, "Models of bus queueing at curbside stops," *Transportation Science*, vol. 49, pp. 204–212, 2014.

[17] A. M. El-Geneidy, J. Horning, and K. J. Krizek, "Analyzing transit service reliability using detailed data from automatic vehicular locator systems," *Journal of Advanced Transportation*, vol. 45, pp. 66–79, 2011.

[18] Y. Zhou, L. Yao, Y. Chen, Y. Gong, and J. Lai, "Bus arrival time calculation model based on smart card data," *Transportation Research Part C: Emerging Technologies*, vol. 74, pp. 81–96, 2017.

[19] Y. Du, F. Deng, F. Liao, and Y. Ji, "Understanding the distribution characteristics of bus speed based on geocoded data," *Transportation Research Part C: Emerging Technologies*, vol. 82, pp. 337–357, 2017.

[20] D. Zhao, W. Wang, A. Woodburn, and M. S. Ryerson, "Isolating high-priority metro and feeder bus transfers using smart card data," *Transportation*, vol. 44, pp. 1535–1554, 2017.

[21] Y. Bie, X. Gong, and Z. Liu, "Time of day intervals partition for bus schedule using GPS data," *Transportation Research Part C: Emerging Technologies*, vol. 60, pp. 443–456, 2015.

[22] M. M. Nesheli and A. Ceder, "Improved reliability of public transportation using real-time transfer synchronization," *Transportation Research Part C: Emerging Technologies*, vol. 60, pp. 525–539, 2015.

[23] C. Ma and X. Xu, "Providing spatial-temporal priority control strategy for BRT lanes: a simulation approach," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, Article ID 04020060, 2020.

[24] N. Saade, J. Doig, and M. J. Cassidy, "Scheduling lane conversions for bus use on city-wide scales and in time-varying congested traffic," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 248–260, 2018.

[25] D. Yang, S. Zheng, C. Wen, P. J. Jin, and B. Ran, "A dynamic lane-changing trajectory planning model for automated vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 228–247, 2018.

[26] A. Alsger, B. Assemi, M. Mesbah, and L. Ferreira, "Validating and improving public transport origin–destination estimation algorithm using smart card fare data," *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 490–506, 2016.

[27] Z. Liu, Y. Liu, Q. Meng, and Q. Cheng, "A tailored machine learning approach for urban transport network flow estimation," *Transportation Research Part C: Emerging Technologies*, vol. 108, pp. 130–150, 2019.

[28] C. Ma and D. Yang, "Public transit network planning in small cites considering safety and convenience," *Advances in Mechanical Engineering*, vol. 12, pp. 1–12, 2020.

[29] X. Ma and Y. Wang, "Development of a data-driven platform for transit performance measures using smart card and GPS

data," *Journal of Transportation Engineering*, vol. 140, Article ID 04014063, 2014.

[30] Y. Yan, Z. Liu, and Y. Bie, "Performance evaluation of bus routes using automatic vehicle location data," *Journal of Transportation Engineering*, vol. 142, Article ID 04016029, 2016.

[31] H. Yu, D. Chen, Z. Wu, X. Ma, and Y. Wang, "Headway-based bus bunching prediction using transit smart card data," *Transportation Research Part C: Emerging Technologies*, vol. 72, pp. 45–59, 2016.

[32] Z. Dai, X. Ma, and X. Chen, "Bus travel time modelling using GPS probe and smart card data: a probabilistic approach considering link travel time and station dwell time," *Journal of Intelligent Transportation Systems*, vol. 23, pp. 175–190, 2019.

[33] M. M. Rahman, S. C. Wirasinghe, and L. Kattan, "Analysis of bus travel time distributions for varying horizons and real-time applications," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 453–466, 2018.

[34] C. R. Rao and Y. Wu, "Linear model selection by cross-validation," *Journal of Statistical Planning and Inference*, vol. 128, pp. 231–240, 2005.

Research Article

Optimal Driving Range for Battery Electric Vehicles Based on Modeling Users' Driving and Charging Behavior

Zhenbo Lu,[1] **Qi Zhang** (ID)**,**[2] **Yu Yuan,**[2] **and Weiping Tong**[2]

[1]*Intelligent Transportation System Research Center, Southeast University, Nanjing, China*
[2]*Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transportation, Southeast University, Nanjing, China*

Correspondence should be addressed to Qi Zhang; zhangqi_1008@seu.edu.cn

This paper proposes a simulation approach for the optimal driving range of battery electric vehicles (BEVs) by modeling the driving and charging behavior. The driving and charging patterns of BEV users are characterized by reconstructing the daily travel chain based on the practical data collected from Shanghai, China. Meanwhile, interdependent behavioral variables for daily trips and each trip are defined in the daily trip chain. To meet the goal of the fitness of driving range, a stochastic simulation framework is established by the Monte Carlo method. Finally, with consideration of user heterogeneity, the optimal driving range under different charging scenarios is analyzed. The findings include the following. (1) The daily trip chain can be reconstructed through the behavioral variables for daily trips and each trip, and there is a correlation between the variables examined by the copula function. (2) Users with different daily travel demand have a different optimal driving range. When choosing a BEV, users are recommended to consider that the daily vehicle kilometers traveled are less than 34% of the battery driving range. (3) Increasing the charging opportunity and charging power is more beneficial to drivers who are characterized by high daily travel demand. (4) On the premise of meeting travel demand, the beneficial effects of increased fast-charging power will gradually decline.

1. Introduction

Battery electric vehicles (BEVs) have the outstanding advantages in zero tailpipe emissions, low noise, convenient maintenance, and high energy conversion efficiency. The deployment of BEVs helps to reduce oil dependence, improve air quality, and reduce pollutions and greenhouse gas emissions [1]. Promoting the development of BEVs is considered as one of the promising solutions for the treatment of severe air pollution in metropolises [2]. The incentives, such as subsidies and tax credits, have effectively promoted the public acceptance for switching to BEVs. For instance, in many mega cities in China, like Beijing, Shanghai, and Hangzhou, the number of vehicle license plates issued per month (car ownership) is under strict control, and the local government has also launched the free licensing policies for BEVs [3].

However, due to the limited battery capacity and charging facilities, inconvenient charging is still an important obstacle to the promotion of BEVs [4]. Compared to the conventional internal combustion engine vehicles (ICEV), BEVs have a shorter driving range, generally 150 km–400 km. Meanwhile, it usually takes hours to charge [5, 6]. Potential customers have repeatedly been found to prefer vehicles with considerably higher available range because of the range anxiety [7]. Although the long driving range design helps to alleviate the user's range anxiety, it results in a higher expenditure on purchase and simultaneously, the affordability and cost-effectiveness is lowered [8]. Optimizing driving range of BEVs based on users' daily travel demand is one of the feasible ways to solve this problem, and it is also the direction of breakthrough for this paper.

The main contribution of this paper is optimizing the driving range for different types of BEV users based on the real-world BEV usage data, specifically including the following: (1) constructing the daily trip chain for BEV drivers

by combining behavioral variables for daily trip and each trip; (2) applying the copula function to examine the dependence between daily travel variables; (3) proposing a simulation framework for the Mote Carlo method to simulate the daily trip chain; and (4) quantifying the fitness of driving range to find the optimal driving range for different drivers.

The remainder of the paper is organized as follows: after a review of relevant literature, Section 3 defines the research problem, followed by the stochastic formulation of driver behavioral models and the procedure of the Monte Carlo simulation in Section 4. Results and discussion are presented in Section 5, and Section 6 concludes this study.

2. Literature Review

Researchers have conducted a series of studies on the travel and charging behaviors of users based on the trajectory data of new energy vehicles (NEVs). In order to understand travel and charging behaviors, many existing studies focused on the distribution of distance traveled per day [5, 9, 10], the start time of charging [11–13], daily vehicle kilometers traveled (DVKT) [10, 12, 14], distance traveled between consecutive charges [5, 15, 16], and battery state of charge (SOC) before and after charging [11, 17, 18]. For example, Wu et al. [9] collected the travel data of 403 plug-in hybrid cars and analyzed the effects of the daily mileage and the mileage between the two charges. Pearre et al. [5] analyzed 470 vehicles with more than 50 days' travel data in Atlanta, USA, and found that the highest probability of DVKT distribution was 19.2–25.6 km; a wide variation in driving distances was observed from different drivers.

Based on the driving behaviors of ICEV and NEV users, how to optimize battery capacity or driving range of electric vehicles (EVs) is studied. Li et al. [19] proposed a hybrid distribution model to describe the daily travel mileage for the purpose of optimizing battery capacity; the testing results indicated that the mixture distribution model could satisfy various drivers. In addition, Dong and Lin [15] proposed the concept of BEV feasibility through a stochastic modeling approach to characterize BEV drivers' behavior. The range of comfort level of the drivers with different driving characteristics was studied to explore solutions to reduce range anxiety. However, although the driving patterns characterizing BEVs are stochastic, the researchers found that the variables are interdependent. After modeling the dependence structure between six variables using a nonparametric copula function, Brady and O'Mahony [20] applied a stochastic simulation methodology to generate a schedule of daily travel and charging profile.

The travel pattern of a particular BEV owner is relatively constant, yet the uncertainties make the travel demand vary from different drivers (i.e., interdriver heterogeneity). Using the data collected from 50 BEVs in Shanghai, China, Yang et al. [11] found the interdriver heterogeneity existing in driving and charging behaviors in terms of distance traveled per day, start time of charging, daily number of charging events, distance traveled between consecutive charges, SOC before and after charging, and time-of-day electricity

demand. Due to the heterogeneities, it is unrealistic for vehicle manufacturers to make the driving range exactly equal to drivers' travel demand. The method of Monte Carlo simulation with the trip chain theory is often used to depict drivers' stochastic behaviors. For example, based on the survey data of National Household Travel Survey (NHTS), Jianfeng et al. [12] fitted the characteristics of the trip chain and analyzed the charging demand by the Monte Carlo simulation. Shuqiang et al. [21] proposed a method to analyze the charging demand of electric vehicles based on the trip chain theory; the Monte Carlo method was applied to explore the probability distribution characteristics of EVs parking time in different regions during one-day travel.

In summary, for the study of optimizing driving range, it is reasonable to consider the correlation between behavior variables and the heterogeneity of users in the modeling process. In addition, practical BEV usage data is rarely utilized to support users daily trip chain reconstruction based on travel and charging behavior. In view of the above methods and shortcomings, this paper proposes a probability distribution model based on the daily user trip chain by considering the correlation between the variables of daily trips and each trip. Combining different scenario settings, users with different driving behavior are analyzed for optimal driving range using Monte Carlo simulation method.

3. Problem Statement

3.1. Daily Trip Chain for BEV Users. The term of "trip chain" has different definitions [22]. To describe each individual's daily movement, we define the trip chain as a sequence of trips bounded by dwells. A trip that commutes between two anchor destinations, such as home to work, or work to home, is powered by electricity, and the dwell between two consecutive trips offers charging opportunities for BEVs.

Assuming BEV drivers leave home in the morning of the travel day with a full battery, the departure time of the first trip T_0 is recorded. With the travel time of the first trip t_1 and the travel speed of the first trip v_1, the distance traveled and the energy consumed for the first trip are estimated. The departure time for the next trip is determined by the dwell time after the first trip s_1. Drivers may plug in their vehicles at the dwell place with considerations that the dwell time is long enough, the SOC is below a certain value (SOC before charging: SOC_n'), and the charging facility is available. After the total number of trips N that day, the driver returned home, and daily vehicle kilometers traveled D are recorded. The electric energy consumed by the n-th trip E_n^t can be estimated by the travel time of the n-th trip t_n and the average travel speed of the n-th trip v_n. The energy obtained from the n-th dwell E_n^c is determined by the time length of the n-th dwell s_n and the power of charging facility. Thus, the daily trip chain is formed by the abovementioned determinants, which include a variety of behavioral variables that describe daily trips and each trip. Table 1 summarizes the behavioral variables of the daily trip chain. The distributions of behavioral variables are derived from the BEV dataset collected from Shanghai, China, in the next section. Once the coverage of charging infrastructure is given, the fitness of

Table 1: A summary of the behavioral variables of the daily trip chain.

Variables for daily trips	Number of daily trips N
	The departure time of the first trip T_0
	Daily vehicle kilometers traveled D
Variables for each trip	The travel time of the nth trip t_n
	The average travel speed of the nth trip v_n
	Dwell time between two consecutive trips s_n
Variables for charging	SOC before charging SOC_n'

the driving range, indicating whether the BEV driving range is capable of covering the daily trip chain, is quantified.

3.2. Correlation between Behavioral Variables.

The behavioral variables in the daily trip chain are mutually influential and not independently distributed. Studies have demonstrated the correlation between the variables for daily trips [20] (i.e., number of daily trips N, the departure time of the first trip T_0, and daily vehicle kilometers traveled D) and applied the copula function to combine the distributions between them. The copula function is called a "join function" or a "dependent function," which is a function that connects the joint distribution of multiple random variables with their respective marginal distributions [23]. Specifically, in accordance with Sklar's theorem [24], the continuous random variables $x_1, x_2, \ldots, x_n$ with cumulative distribution functions (CDFs) $F_1(x_1), F_2(x_2), \ldots, F_n(x_n)$, respectively, are joined by copula C if their joint distribution function can be expressed as

$$F(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)). \quad (1)$$

Further, the three behavior variables of each trip not only affect each other, but also depend on the three variables of daily trip. Equation (2) establishes the link between each trip variable and the daily trip variable:

$$D = \sum_n^N t_n \cdot v_n. \quad (2)$$

Daily vehicle kilometers traveled D are the cumulative sum of the travel time of each trip t_n multiplied by the average speed of each trip v_n under the number of daily trips N.

3.3. Fitness of Driving Range.

Due to limited battery range and insufficient charging infrastructure, the travel demand of BEV users is likely under constraint. BEV users may check the remaining SOC before each travel. Once the travel distance is beyond the remaining range, drivers have to charge the battery or change the travel plan. This situation is referred to as "range limitation" [15]. The more times the mileage limit occurs, the less user satisfaction will be.

We reconstruct the daily trip chain for BEV users by simulation. The remaining SOC is determined by the energy consumed in the previous travels and the availability of charging opportunity. Once there is no charging opportunity, the situation of "range limitation" may occur when the

remaining battery range is not long enough for the next trip. The trip chain must be terminated as the battery may run out. We count the number of the days with incomplete trip chain and denote it as $\tau(R)$, where R represents driving range. It can be speculated that the larger the battery, the less possibility that the range limitation occurs.

The fitness of driving range θ is defined as

$$\theta = \left(1 - \frac{\tau(R)}{M} \times 100\%\right), \quad (3)$$

where M is the number of travel days for BEV users. The proposed battery fitness measure can incorporate stochastic driver behaviors by incorporating randomly distributed variables in daily trip chain. If the simulated travel days for a BEV are 10,000, the fitness $\theta = 95\%$ means that there are 50 days when the BEV cannot fulfill the traveler's travel demand.

4. Methodology

4.1. Modeling of Driver's Behaviors

4.1.1. Data Description.

This study makes use of a rich database collected from 50 BEVs over a period of 4–12 months (spanning from June 5, 2015, to June 30, 2016). The dataset is provided by Shanghai Electric Vehicles Data Center (SHEVDC) that is developed to remotely monitor electric vehicles driven across the city. The 50 BEVs, used as personal vehicles, are with the same model of Roewe E50, which is a pure electric passenger car with a 22.4 kW·h battery pack and a claimed driving range of 170 km under NEDC conditions [25].

Vehicle terminals, such as global positioning system data loggers and instruments to measure voltage and current, are installed on BEVs for data collection [26, 27]. The collected data includes turn on time, turn off time, total mileage, SOC, voltage, and current. Vehicles' driving information includes time-stamped location (i.e., longitude and latitude), spot speed, and azimuth. After data cleaning and consistency check procedure to remove invalid data and possible errors from the record, BEV owners' driving patterns, such as the trip distance, average travel speed, departure time of each trip, trip time length, SOC before/after each trip, and energy consumed, are extracted.

After the data cleaning of the original dataset, 12,855 valid trips, 7,112 accurate charges, and 7,275 travel days with authentic trips were extracted. Due to the intermittent record, the travel distance between two consecutive charges got 6,244 records. Table 2 summarizes the descriptive statistics of the valid trips and charges.

In terms of travel demand, the indicator that best reflects daily travel demand is the daily travel distance. Its average is 51.6 km, which is only 30% of the battery's driving range. In addition, the average number of daily trips is 2.97, and the highest value is 9, which can be roughly understood as a daily commute user. For charging habits, the average SOC before charging is 45.6%. This means that the user prefers to start connecting the charging plug when there is still about half of SOC remaining.

	Number of samples	Avg.	Med.	Max.	Min.	St.D.
Number of daily trips N	7275	2.97	2	9	1	1.16
The departure time for the first trip T_0	7275	11:23	10:45	23:59	0:00	4:50
Daily vehicle kilometers traveled D (km)	7275	51.6	55.4	95.3	17.5	18.8
Travel time t_n (h)	12855	0.71	0.57	9.12	0.05	0.54
Travel speed v_n (km/h)	12855	19.62	17.04	80	0.25	11.42
Dwell time between two consecutive trips s_n (h)	6244	2.7	1.69	22.15	0.06	2.86
Charging power	7112	3.80	3.21	6.71	1.31	2.64
SOC before charging SOC'_n (%)	7112	45.6	45.4	61.6	18.5	8.3

4.1.2. Modeling Variables for Daily Trips. In order to verify if there is dependence among the variables for daily trips, Pearson's linear correlation which measures a monotonic relationship between variables is estimated and placed in the matrix, R [28]. Pearson's linear correlation derived for the variables representing daily trips: the departure time for the first trip T_0, daily vehicle kilometers traveled D, and number of daily trips N are given by

$$R = \begin{bmatrix} 1 & -0.23 & 0.51 \\ -0.23 & 1 & -0.35 \\ 0.51 & -0.35 & 1 \end{bmatrix}. \quad (4)$$

Testing the hypothesis of no correlation against the alternative that there is a nonzero correlation was applied to the elements of the correlation matrices [29]. All coefficients in the R matrix are found to be statistically significant assuming a 0.05 level of significance. In the R matrix, there is strong evidence that all variables are correlated and it is important to model the dependence structure between the core variables. According to R (1, 2), daily travel distance is negatively correlated with the first departure time of the day, which means that the later the user travels the earliest day, the shorter the daily travel mileage is. Meanwhile, R (1, 3) shows that the more daily trips, the greater the distance traveled each day. Therefore, after demonstrating the dependencies between the simulated variables, a joint PDF using their respective marginal distributions and a copula function needs to be constructed.

By modeling using copula function, we get the joint distribution result between the three variables for daily trips. If the resulting modeling structure is correct, the original and simulated values of the corresponding variables should have the same distribution. This is studied using the quantile-quantile (Q-Q) plot of the simulated and raw data. Figure 1 shows the Q-Q plot of the raw and simulated values of the variables (daily vehicle kilometers traveled). If the two distributions to be compared are similar, the points in the Q-Q diagram will approximate a straight line. It can be seen from the figure that it approximates a straight line, indicating that the simulated and original data are from the same distribution.

4.1.3. Modeling Variables for Each Trip

(1) Travel Time t_n. Figure 2 plots the distribution of travel time for each trip. 54.28% of the travel time is less than 40 minutes. We find that the travel time t_n follows the log-normal distribution [22]. The PDF of t_n is written as
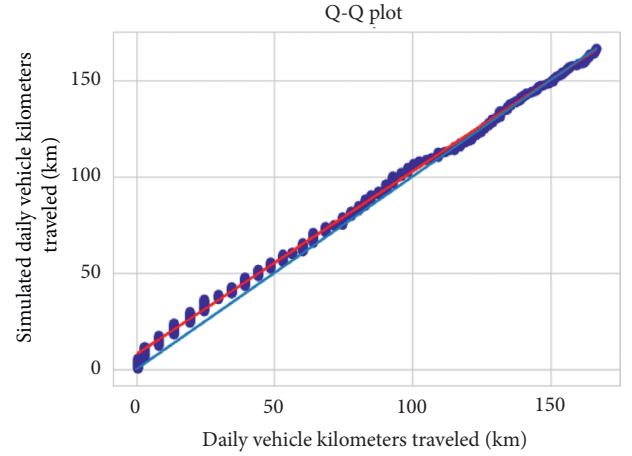

FIGURE 1: Q-Q plot of the simulated and original variable (daily vehicle kilometers traveled).

$$f(t_n) = \frac{1}{\sqrt{2\pi}\sigma t_n} \exp\left[-\frac{1}{2\sigma^2}(\ln t_n - \mu)^2\right], \quad t_n > 0, \quad (5)$$

where $\mu = 2.9115$ and $\sigma = 0.876$.

(2) Travel Speed v_n. The travel speed for each trip follows the beta distribution, with an average value of 22.04 km/h (Figure 3). The PDF of travel speed v_n is expressed as

$$p_3(y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) + \Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}.$$

$$y = \frac{v_n - a}{b}, \quad (6)$$

where $\alpha = 2.9339$, $\beta = 3256600$, $a = 0.50467$, and $b = 2399300$.

(3) Dwell Time between Two Consecutive Trips s_n. Once a trip is ended, BEV drivers may stop at the destination for a certain period. The dwell time is one of the key factors for charging decision [30]. Figure 4 depicts the distribution of dwell time between two consecutive trips. 57.5% of the dwells last for less than 2 hours, which may not be appropriate for charging with slow charging. The dwell time is estimated by a lognormal distribution. The PDF of s_n is written as

$$p_4(s_n) = \frac{1}{\sqrt{2\pi}\sigma s_n} \exp\left[-\frac{1}{2\sigma^2}(\ln s_n - \mu)^2\right], \quad x > 0, \quad (7)$$

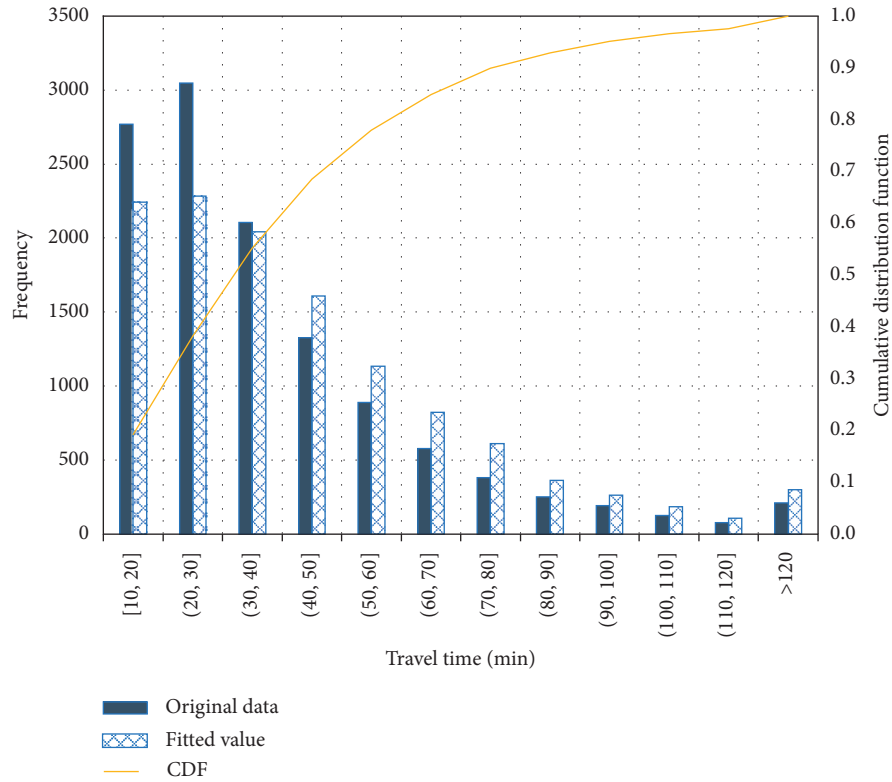where $\mu = 2.6$ and $\sigma = 1.0147$.
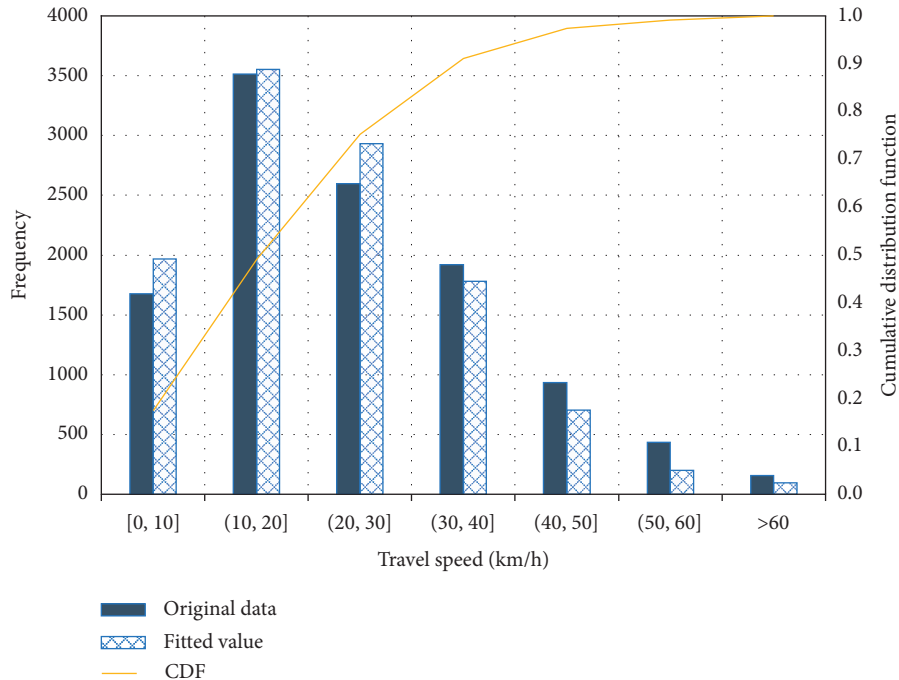
FIGURE 2: The distribution of travel time.

FIGURE 3: The distribution of travel speed.

(4) SOC before Charging, SOC_n'. The SOC before charging is adopted as the threshold for each charging event. Once the remaining SOC is above the threshold, the charging event will not occur [31]. As shown in Figure 5, the preference of the SOC before charging is not obvious, indicating drivers are likely to charge their vehicles whenever they have charging opportunities. The distribution of SOC_n' is found subject to the Johnson SB distribution. The PDF of SOC_n' is written as
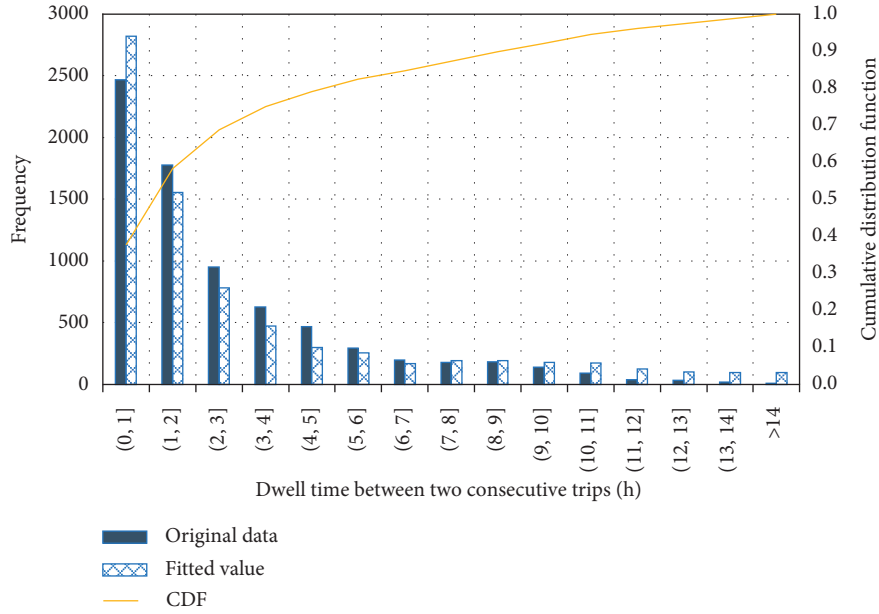
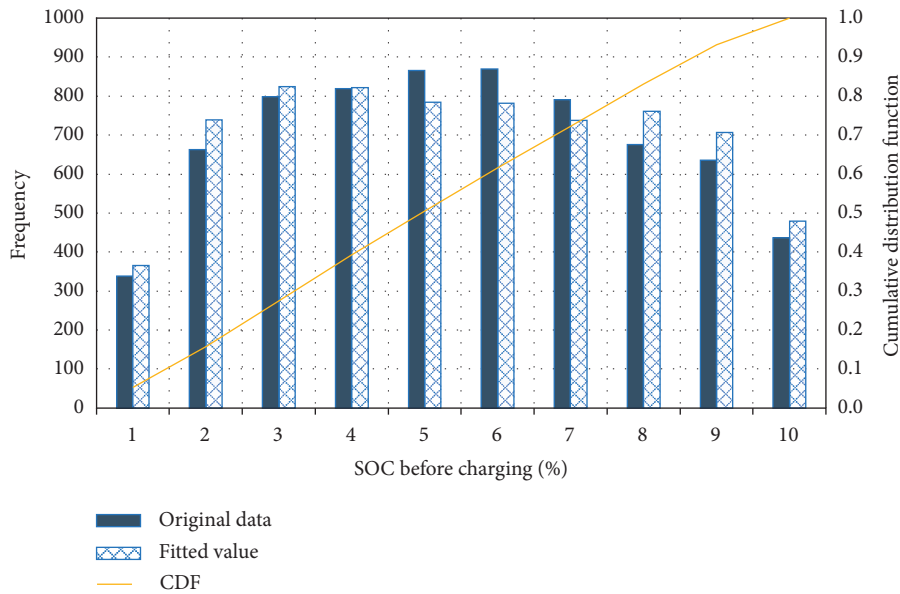FIGURE 4: The distribution of dwell time between two consecutive trips.

FIGURE 5: The distribution of SOC before charging.

$$p_5(t_n) = \frac{\delta}{\sqrt{2\pi}z(1-z)}\exp\left(-\frac{1}{2}\left(\gamma + \delta\ln\left(\frac{z}{1-z}\right)\right)^2\right), \quad \zeta \leq z \leq \zeta + \lambda,$$

$$z = \frac{SOC_n' - \lambda}{\lambda}$$

$$(8)$$

where $\gamma = 0.081$, $\delta = 0.242$, $\zeta = 0$, and $\lambda = 100$.

4.1.4. Drivers Heterogeneity. In order to consider the impact of heterogeneity, Yang et al. [11] applied a machine learning method to analyze BEV drivers' habitual behaviors based on

the same sample data. User heterogeneity is mainly reflected in different travel patterns. Daily vehicle kilometers traveled D are not only the core of the behavior variable for daily trips, but also the dependency of each trip variables. Therefore, the research on user heterogeneity in this paper mainly focuses on the optimal driving range of different daily travel users. According to the results of Yang et al. [11], the 50 BEV sample vehicles can be divided into 4 clusters in Table 3 by daily vehicle kilometers traveled D. Obviously, based on the same charging scenario restrictions, users with different travel pattern will have corresponding optimal driving range.

TABLE 3: Summary metrics of different clusters.

Type	Number of vehicles	Avg. daily vehicle kilometers traveled (km)		
		Sample size	Mean	St.D.
Cluster A	34	4406	58.0	42.0
Cluster B	2	240	93.4	44.3
Cluster C	5	1059	31.7	30.1
Cluster D	9	1570	23.3	19.0

4.2. Monte Carlo Simulation

4.2.1. Simulation Assumptions. The data set of this study is based on the Roewe E50, so the initial driving range R during the simulation is also 170 km. That is, the driving range limitation of this study is specifically defined for this vehicle model. Several conditions are set for the simulation context/environment. They are summarized as follows.

Condition 1. Charging places: two scenarios are set for charging locations. In Scenario 1, users can only charge the battery at home when they go back home after the last trip of the travel day. In Scenario 2, the charging facility is available at workplace, and users can choose to charge at the workplace or charge at home. Two time periods for workplace charging are set, i.e., from 09 : 00 to 16 : 00 [32].

Condition 2. Energy consumption: the electricity consumed for the nth trip E_n^t (kW) is supposed to be determined by the travel time and travel distance [33]. A linear relationship is found based on the real-world data with the R-squared value of 0.937. It can be written as

$$E_n^t = 0.0038t_n + 0.8358l_n + 0.9829, \qquad (9)$$

where l_n (km) is travel distance for the nth trip determined by v_n and t_n.

Condition 3. Charged energy: the amount of electricity charged mainly depends on the charging power and the dwell time. The charging power is assumed as constant, and the charged electricity E_n^c (kW) for the nth dwell can be written as

$$E_n^c = \min\left(R_n^r + p \times s_n, R\right), \qquad (10)$$

where R_n^r (kW·h) is the remaining range after the nth trip, p is the charging power, and R (kW·h) is the battery capacity.

4.2.2. Simulation Process. Based on the distribution models of drivers' driving and charging behaviors, the Monte Carlo simulation is applied to simulate the trip chain for BEV users. To stabilize the simulation results, the number of the simulated travel days M is set as 10,000. The simulation process is presented in Algorithm 1.

5. Results and Discussion

The charging problem is a key obstacle hindering the development of BEVs. The feasible methods to solve the charging problem mainly include increasing the charging opportunity and increasing the power of the charging pile. These two feasible methods will be studied to influence the optimal driving range of different users in this section.

5.1. Impact of Workplace Charging. In general, household charging piles mainly apply alternating current (AC), which is mainly divided into two levels. Level 1 operates at 120 VAC (voltage in an alternating current), while Level 2 uses 208 or 240 VAC, and the corresponding charging power ranges from 1.4 kW to 7.2 kW. If the charging pile in public places uses direct current (DC), the charging power can reach 50 kW or more. Although the charging power of Electric Vehicle Supply Equipment (EVSE) is now getting bigger, even more than 100 kW, it takes a long time to spread to the average drivers and to apply to each BEV model [34,35]. Almost all charging events extracted from the dataset are performed using AC Level 2 Electric Vehicle Supply Equipment (EVSE), with an average value of 3.8 kW in Table 2. Thus, by setting three kinds of slow charging power (i.e., the power of 3.5 kW, 7 kW, and 15 kW) at two charging places, the impact of increasing the charging opportunity on the optimal driving range is studied. Given the fitness requirement $\theta = 95\%$, Figure 6 presents the results of optimal driving range in Scenario 1 and Scenario 2.

Whether it is to increase the charging opportunity or the charging power, the value of the optimal driving range can be reduced, because the daily travel demand of users is more easily met. In detail, the daily vehicle kilometers traveled of clusters A and B exceed the average value in Table 2. Compared with Scenario 1, Scenario 2 that increases the charging opportunity at the workplace has a significant effect on reducing the optimal driving range, especially for cluster B with a daily vehicle kilometers traveled value of 93 km. Clusters C and D with smaller travel demand are not sensitive enough to Scenario 1 and 2 with different charging opportunities. When the charging power is 15 kW, the optimal driving range of both clusters is only slightly reduced.

According to the original driving range of the Roewe E50 (170 km), two drivers in cluster B are not suitable for this model unless the charging power and charging opportunities are appropriately increased. Also, the daily travel demand of the 14 users in clusters C and D can be mostly satisfied even if charging only at home and the charging power is 3.5 kW. As the daily vehicle kilometers traveled of cluster A are similar to the average value, it can be found that when the average daily travel distance of the user accounts for about 30%–34% of the battery driving range, the fitness of battery capacity cannot be satisfied. Therefore, users should choose their own BEVs whose daily travel demand is within 34% of the battery driving range.

5.2. Impacts of Fast Charging. In order to understand the impact of fast charging, based on the results of Section 5.1, the benchmark scenario is set to 15 kW at home and at the workplace. Assuming fast charging is only available at workplace, the added charging power of workplaces (CPW)

Input: Practical data of the behavioral variables of the daily trip chain in Table 1
Output: Simulated travel and charging process
Process:
Initial: $M = 10{,}000$, $m = 1$, $i = 0$.
   Step 0: the vehicle is assumed to depart from home with a full battery
   Step 1: generate the daily trip, the number of daily trips N, the departure time for the first trip T_0, and the daily travel distance D using copula function; set $i = 1$
   Step 2: generate the travel distance, the average travel speed for each trip, and the dwell time between the two trips by equation (2) on the day m
   Step 3: calculate the electricity consumed for the ith trip; check whether the remaining SOC is sufficient for the ith trip. If it is, update the remaining SOC after arriving at the destination, and proceed to Step 4; otherwise, the ith trip is marked as failed with "range limitation," and choose alternative travel mode to make up for this trip; set $m = m+1$, and return to Step 0
   Step 4: generate SOC before charging SOC_n'; check whether a charging opportunity is available at that time. If the charge is performed, calculate the increased SOC and update the remaining range; otherwise, the remaining range keeps the same
   Step 5: check whether i equals N. If it does, move to Step 6; otherwise, $i = i + 1$, and return to Step 2
   Step 6: check whether m equals M. If it does, the simulation is terminated; otherwise, set $m = m + 1$, and return to Step 1

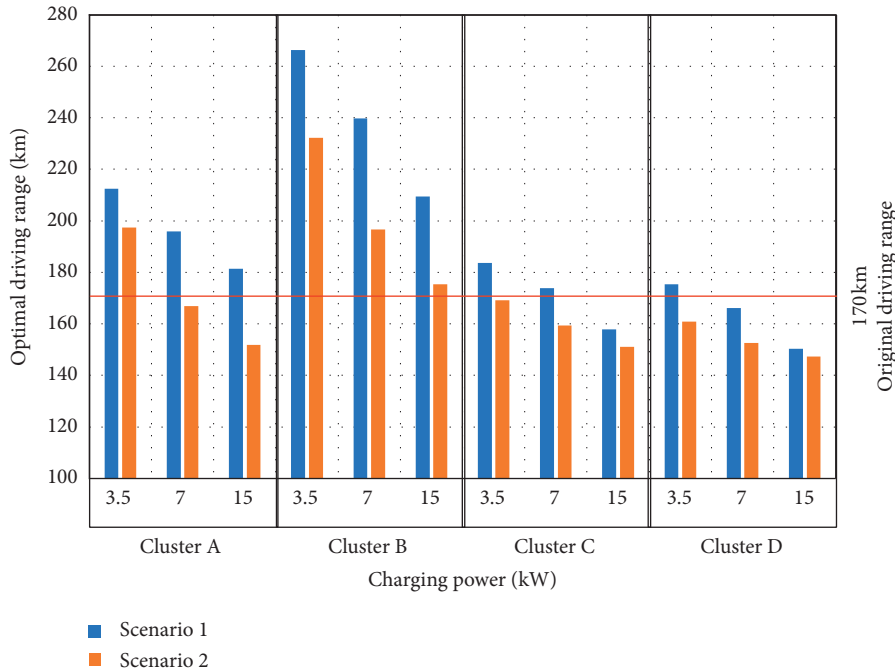ALGORITHM 1: Summary of simulation methodology.

FIGURE 6: Optimal driving range for two scenarios when $\theta = 95\%$.

is set to 30 kW, 45 kW, and 50 kW, respectively. The charging power at home (CPH) is still set to 15 kW. Figure 7 shows a stacked bar chart of relative reduction percentages when the optimal driving range is compared with the values of the benchmark scenario as CPW increases.

Cluster B, which has the largest daily vehicle kilometers traveled, has the largest reduction percentage. When CPW = 50 kW and CPH = 15 kW, the optimal driving range can be reduced by 10.4% compared with the benchmark scenario. Conversely, the cumulative reduction percentage of clusters C and D is in the range of 7.6%–8.4%. The benefits of fast charging are more significant to users with high daily travel demand. Besides, as the CPW increases,

the value of the relative reduction percentage for each cluster gradually decreases. It can be seen that, on the premise of meeting travel demand, the beneficial effects brought by the increase in fast-charging power will gradually weaken.

To sum up, with the increase in daily travel demand of residents, the beneficial effects of charging piles equipped with fast-charging power can be reflected. The government should pay more attention to the layout of the fast-charging piles in the area where residents with large daily travel needs are located. Simultaneously, a reasonable amount of fast-charging power should be studied without the need for blind high power.
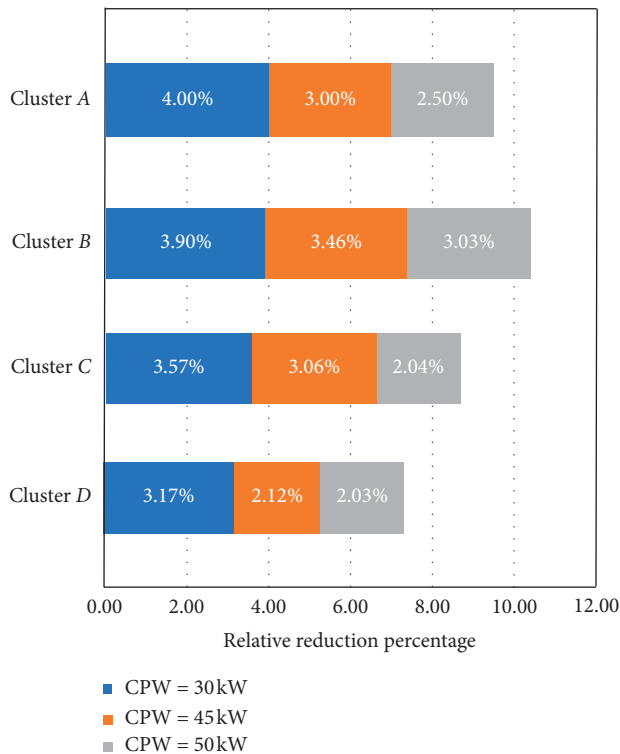
FIGURE 7: A stacked bar chart of relative reduction percentages.

6. Conclusions

Aimed at finding out the optimal driving range considering user heterogeneity in different charging scenarios, this study proposes a simulation modeling method representing BEV users' driving and charging behavior using the real-world in-use data collected from Shanghai, China. The Monte Carlo simulation is adopted to reconstruct BEV users' daily trip chains and quantify the fitness of the driving range. The key findings from the results include the following. (1) The daily travel chain can be reconstructed by variables for daily trips and each trip, and dependencies exist in these variables. (2) According to the results of four clusters, users with different daily travel demand have different optimal driving ranges. When choosing their BEV, users should consider that the daily vehicle kilometers traveled are less than 34% of the battery driving range. (3) Increasing both the charging opportunity and charging power is more beneficial to drivers with large daily vehicle kilometers traveled. (4) With the increase of fast-charging power, this beneficial effect gradually weakens when daily travel demand is met.

The main contribution of this paper is to construct a daily trip chain to analyze the optimal driving ranges of various heterogeneous users. However, due to errors in the vehicle's latitude and longitude in the data source, the restriction is that only the home and workplace are assumed during the daily trip chain. The purpose of daily activities of residents is diverse, not only work and home. Therefore, further enriching the daily trip chain is also the main focus in the future research direction.

References

[1] R. R. Desai, R. B. Chen, and W. Armington, "A pattern analysis of daily electric vehicle charging profiles: operational efficiency and environmental impacts," *Journal of Advanced Transportation*, vol. 2018, Article ID 6930932, 15 pages, 2018.

[2] A. D. Alvarez, F. S. Garcia, J. E. Naranjo, J. J. Anaya, and F. Jimenez, "Modeling the driving behavior of electric vehicles using smartphones and neural networks," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 3, pp. 44–53, 2014.

[3] Y. Guo, Z. Li, Y. Wu, and C. Xu, "Evaluating factors affecting electric bike users' registration of license plate in China using bayesian approach," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 59, pp. 212–221, 2018.

[4] K. Hu, J. Wu, and T. Schwanen, "Differences in energy consumption in electric vehicles: an exploratory real-world study in Beijing," *Journal of Advanced Transportation*, vol. 2017, Article ID 4695975, 17 pages, 2017.

[5] N. S. Pearre, W. Kempton, R. L. Guensler, and V. V. Elango, "Electric vehicles: how much range is required for a day's driving?" *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1171–1184, 2011.

[6] M. Helmbrecht, C. Olaverri-Monreal, K. Bengler, R. Vilimek, and A. Keinath, "How electric vehicles affect driving behavioral patterns," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 3, pp. 22–32, 2014.

[7] A. Dimitropoulos, P. Rietveld, and J. N. V. Ommeren, "Consumer valuation of changes in driving range: a meta-analysis," *Transportation Research Part A: Policy and Practice*, vol. 55, pp. 27–45, 2011.

[8] T. Miwa, H. Sato, and T. Morikawa, "Range and battery depletion concerns with electric vehicles," *Journal of Advanced Transportation*, vol. 2017, Article ID 7491234, 12 pages, 2017.

[9] X. Wu, D. Freese, A. Cabrera, and W. A. Kitch, "Electric vehicles' energy consumption measurement and estimation," *Transportation Research Part D: Transport and Environment*, vol. 34, no. 34, pp. 52–67, 2015.

[10] Z. Lin, J. Dong, C. Liu, and D. greene, "Estimation of energy use by plug-in hybrid electric vehicles," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2287, no. 1, pp. 37–43, 2012.

[11] J. Yang, J. Dong, Q. Zhang, Z. Liu, and W. Wang, "An investigation of battery electric vehicle driving and charging

behaviors using vehicle usage data collected in Shanghai, China," *Transportation Research Record Journal of the Transportation Research Board*, vol. 2672, 2018.

[12] W. Jianfeng, T. Shun, X. Xiangning, L. Chen, and L. Kunyu, "Analysis on charging demand of ev based on stochastic simulation of trip chain," *Power System Technology*, vol. 39, no. No. 6, pp. 1477–1485, 2015.

[13] Y. Liu, Z. Liu, and R. Jia, "DeepPF: a deep learning based architecture for metro passenger flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 18–34, 2019.

[14] T. Franke and J. F. Krems, "What drives range preferences in electric vehicle users?" *Transport Policy*, vol. 30, no. 3, pp. 56–62, 2013.

[15] J. Dong and Z. Lin, "Stochastic modeling of battery electric vehicle driver behavior," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2454, no. 1, pp. 61–67, 2014.

[16] Q. Cheng, Z. Liu, and W. Y. Szeto, "A cell-based dynamic congestion pricing scheme considering travel distance and time delay," *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, pp. 1286–1304, 2019.

[17] J. Bao, P. Liu, X. Qin, and H. Zhou, "Understanding the effects of trip patterns on spatially aggregated crashes with large-scale taxi GPS data," *Accident Analysis & Prevention*, vol. 120, pp. 281–294, 2018.

[18] A. S. Al-Ogaili, T. J. Tengku Hashim, N. A. Rahmat et al., "Review on scheduling, clustering, and forecasting strategies for controlling electric vehicle charging: challenges and recommendations," *IEEE Access*, vol. 7, pp. 128353–128371, 2019.

[19] Z. Li, S. Jiang, J. Dong, S. Wang, Z. Ming, and L. Li, "Battery capacity design for electric vehicles considering the diversity of daily vehicles miles traveled," *Transportation Research Part C: Emerging Technologies*, vol. 72, pp. 272–282, 2016.

[20] J. Brady and M. O'Mahony, "Modelling charging profiles of electric vehicles based on real-world electric vehicle charging data," *Sustainable Cities and Society*, vol. 26, pp. 203–216, 2016.

[21] Z. Shuqiang, Z. Jingren, L. Zhiwei, and Z. Shuo, "EV charging demand analysis based on trip chain theory," *Electric Power Automation Equipment*, vol. 37, no. 8, pp. 105–117, 2017.

[22] E. Chen, Z. Ye, C. Wang, and M. Xu, "Subway passenger flow prediction for special events using smart card data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, 2019.

[23] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 40–49, 2008.

[24] M. Sklar, "Fonctions de repartition an dimensions et leurs marges," *Publications from the Institute of Statistics of the University of Paris*, vol. 8, pp. 229–231, 1959.

[25] ROEWE, *The Introcution of Roewe E50*, ROEWE, Shanghai, China, 2018, http://www.roewe.com.cn/roewee50/.

[26] X. Gu, M. Abdel-Aty, Q. Xiang, Q. Cai, and J. Yuan, "Utilizing UAV video data for in-depth analysis of drivers' crash risk at interchange merging areas," *Accident Analysis & Prevention*, vol. 123, pp. 159–169, 2019.

[27] Y. Yuan, M. Yang, J. Wu, S. Rasouli, and D. Lei, "Assessing bus transit service from the perspective of elderly passengers in Harbin, China," *International Journal of Sustainable Transportation*, vol. 13, pp. 1–16, 2019.

[28] C. Wang, C. Xu, J. Xia, Z. Qian, and L. J. T. r. p. C. e. t. Lu, "A combined use of microscopic traffic simulation and extreme value methods for traffic safety evaluation," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 281–291, 2018.

[29] D. Kurowicka and R. M. Cooke, *Uncertainty Analysis with High Dimensional Dependence Modelling*, John Wiley & Sons, Hoboken, NJ, USA, 2006.

[30] L. Li, J. Zhang, Y. Wang, and B. Ran, "Missing value imputation for traffic-related time series data based on a multi-view learning method," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, 2018.

[31] T. Franke, M. Günther, M. Trantow, and J. F. Krems, "Does this range suit me? Range satisfaction of battery electric vehicle users," *Applied Ergonomics*, vol. 65, pp. 191–199, 2017.

[32] V. Torres-Sanz, J. A. Sanguesa, F. J. Martinez, P. Garrido, and J. M. Marquez-Barja, "Enhancing the charging process of electric vehicles at residential homes," *IEEE Access*, vol. 6, pp. 22875–22888, 2018.

[33] Y. Pan, S. Chen, F. Qiao, S. V. Ukkusuri, and K. Tang, "Estimation of real-driving emissions for buses fueled with liquefied natural gas based on gradient boosted regression trees," *Science of the Total Environment*, vol. 660, pp. 741–750, 2019.

[34] H. Wang, D. Zhao, Q. Meng, G. P. Ong, and D.-H. Lee, "A four-step method for electric-vehicle charging facility deployment in a dense city: an empirical study in Singapore," *Transportation Research Part A: Policy and Practice*, vol. 119, pp. 224–237, 2019.

[35] A. Dubey and S. Santoso, "Electric vehicle charging on residential distribution systems: impacts and mitigations," *IEEE Access*, vol. 3, pp. 1871–1893, 2015.

WILEY | Hindawi

Research Article

Price Regulation Mechanism of Travelers' Travel Mode Choice in the Driverless Transportation Network

Jianhui Wu,[1] **Yuanfa Ji,**[2] **Xiyan Sun** ⓘ**,**[2] **and Yan Xu**[1]

[1]*School of Information Science and Technology, Hunan Institute of Science and Technology, Yueyang 414006, China*
[2]*Guangxi Key Laboratory of Precision Navigation Technology and Application, Guilin University of Electronic Technology, Guilin 541004, China*

Correspondence should be addressed to Xiyan Sun; sunxiyan1@163.com

To study the guidance method of driverless travel mode choice from the perspective of traffic supply-demand, we assume that all vehicles are driverless and establish a multimodal travel market model to depict the supply-demand relationship of multimodal driverless transportation network. To regulate the disequilibrium multimodal travel market, an optimal price regulation law is proposed, which aims to minimize the supply-demand deviation and the amplitude of price regulation. Then, the existence, uniqueness, and stability of the optimal price regulation law are confirmed. In the calculation process of a numerical example, the travel prices of driverless car and driverless subway are realized by congestion fee and subway fare, respectively. The results indicate that the optimal price regulation law can reduce the supply-demand deviation of the multimodal travel market and guide travelers to choose a reasonable travel mode to travel in the driverless transportation network.

1. Introduction

With the development of driverless technology, the city is again at the crossroads of a historic transformation in the transportation technology. Hence, it is important to determine how the impact of driverless technology on urban transportation should be addressed. Such problems have attracted considerable attention from scholars. Morando et al. [1] believed that the high penetration rate of autonomous vehicles can significantly improve the traffic safety in urban transportation network. Ye and Yamamoto [2] pointed out that the advantages of setting dedicated lanes for automatic vehicles are significant when automatic vehicles are in the medium density range of mixed traffic flow. Smolnicki and Sołtys [3] investigated the effects of different driverless mobility solutions on urban spatial structures. Zhang et al. [4] studied the impact of a shared autonomous vehicles system on urban parking demand. Kamel et al. [5] analyzed the effects of user preferences on the modal split of shared autonomous vehicles. Yi et al. [6] studied the effect of the ambient temperature on the energy cost and charging demand for autonomous electric vehicles. Moreno et al. [7] believed that ride sharing can more effectively reduce extra vehicle kilometers, while the same amount of trips by shared autonomous vehicles will require longer empty trips.

Currently, research on urban traffic planning and management in the driverless environment focuses primarily on road planning, parking planning, public transport planning, and traffic management planning. The purpose of road planning is to solve the driving problem of driverless vehicles on the road. Relevant achievements include lane planning (e.g., Liu and Song [8] and and Xia et al. [9]), traffic signal control planning (e.g., Domínguez and Sanguino [10] and Jiang [11]), and speed limit planning (e.g., Tajalli and Hajbabaie [12] and

Liu et al. [13]). The purpose of parking planning is to solve the parking problem of driverless vehicles, and relevant achievements include parking lot design (e.g., Nourinejad et al. [14] and Estepa et al. [15]) and parking lot management (e.g., Yamashita and Takami [16] and Wang et al. [17]). The purpose of public transport planning is to attract some travelers to travel by the large capacity and high occupancy of driverless public transport. Relevant achievements include driverless public transport type planning (e.g., Leich and Bischoff [18] and Abe [19]) and driverless public transport line design (e.g., Cao and Ceder [20] and Tong et al. [21]). The purpose of traffic management planning is to maintain traffic order and ensure smooth traffic and safe run. Relevant achievements include vehicles' safety performance (e.g., Ye and Yamamoto [22] and Virdi et al. [23]), traffic laws and regulations (e.g., Prakken [24] and Bartolini et al. [25]), and traffic control (e.g., Wagner [26] and Tettamanti et al. [27]).

Overall, the new driverless technology has brought great changes to urban transportation system. However, the driverless transportation system is large and complex. There may be congestion on some roads during the peak period, and it may be difficult to find a parking place in office and residential areas during the nonpeak period. Kitamura et al. [28] pointed out that traffic planners need to find effective ways to induce or suppress travel demand in the next generation of transportation planning methodologies. Brideges [29] believed that the "multitraveler sharing" public transport system can solve the problem of urban congestion and proposed a "carrot and stick" mechanism to encourage travelers to share their vehicles with others. Therefore, it is also necessary to strengthen the guidance of travelers' travel mode choice in the driverless transportation network.

Ye and Wang [30] proposed a bilevel programming model of congestion pricing for two travel modes: driverless vehicle and conventional vehicle. Zhang et al. [31] studied the integrated morning-evening commuting pattern at the system optimum in the fully autonomous vehicle environment. On the whole, there are few researchers to study the guidance method of driverless travel mode choice from the relationship of traffic supply-demand. However, the disequilibrium of traffic supply-demand is the fundamental reason for urban traffic congestion. Based on this, this study regards each type of driverless travel modes on each OD pair as a travel market, regards driving time, driving cost, and perceived traffic service quality as travel price, and attempts to seek the price regulation mechanism of multidriverless travel mode in terms of market supply-demand.

The remainder of this paper is organized as follows. In Section 2, we establish a multimodal travel market model. In Section 3, we present an optimal price regulation law of disequilibrium multimodal travel market and prove the existence, uniqueness, and stability of the optimal

price regulation law. A numerical example is demonstrated in Section 4, and the conclusions are presented in Section 5.

2. Multimodal Travel Market Model

To depict the supply-demand relationship of the multimodal driverless transportation network, we developed a multimodal travel market model. In the driverless transportation network, we suppose that all travelers do not have a private car, and they can only travel by driverless vehicles. Furthermore, we assume that there are M types of driverless travel modes (e.g., driverless car, driverless bus, and driverless subway) on OD pair (i, j). If we regard each type of driverless travel mode as a travel market, then there are M types of travel markets on OD pair (i, j).

2.1. Notional Supply and Demand. We define all travelers of the mth travel mode who need to travel from the origin node i to the destination node j as the demand side of the mth travel market. According to the definition of the demand function in the elastic demand traffic assignment problem, we define the linear combination between exogenous variable vector X_{ij}^m and travel price P_{ij}^m as the notional demand $\widetilde{D}_{ij}^m$ of the mth travel market on OD pair (i, j) at the kth period. It can be expressed as

$$\widetilde{D}_{ij}^m(k) = \alpha_{ij}^m \left\{ X_{ij}^m(k) \right\}^T + \beta_{ij}^m P_{ij}^m(k), \\ \forall \beta_{ij}^m < 0, m \in M_{ij}, i, j \in N, i \neq j, \tag{1}$$

where α_{ij}^m represents the influence coefficient vector of X_{ij}^m on $\widetilde{D}_{ij}^m$, β_{ij}^m is the influence coefficient of P_{ij}^m on $\widetilde{D}_{ij}^m$, X_{ij}^m reflects the total travel demand, as well as the traveler's gender, age, occupation, income, etc., $\left\{ X_{ij}^m(k) \right\}^T$ is the transpose of the vector X_{ij}^m, P_{ij}^m is a comprehensive index that reflects the driving time, driving cost, perceived traffic service quality, etc., M_{ij} is the set of all travel modes on OD pair (i, j), and N is the set of all nodes.

We define the transport managers who can provide transport services to the demand-side of the mth travel market on OD pair (i, j) in the short period as the supply side of the mth travel market. Considering that the supply of driverless car (or driverless bus, driverless subway, etc.) in the short period is affected a little by travel price, we ignore the influence of travel price. Suppose that the notional supply $\widetilde{S}_{ij}^m$ of the mth travel market on OD pair (i, j) at the kth period is only related to the exogenous variable vector Y_{ij}^m. It can be expressed as

$$\widetilde{S}_{ij}^m(k) = \eta_{ij}^m \left\{ Y_{ij}^m(k) \right\}^T, \quad \forall m \in M_{ij}, i, j \in N, i \neq j, \tag{2}$$

where η_{ij}^m represents the influence coefficient vector of Y_{ij}^m on $\widetilde{S}_{ij}^m$ and Y_{ij}^m reflects the total supply, severe weather, natural disaster, traffic accident, etc.

2.2. Effective Supply and Demand. Suppose that the supply side between different travel modes is independent of each other, and the excessive demand of one travel market may spill over the effective demand of other travel market in the multimodal travel market. For example, if the demand of the driverless car travel market exceeds the supply, some driverless car travelers may switch to other driverless travel modes, such as driverless bus or driverless subway. Referring to the relevant research results in the field of economics (Browne [32]), we define the linear combination between the notional demand of the mth travel market and the effective supply-demand difference of other travel markets as the effective demand D_{ij}^m of the mth travel market on OD pair (i, j) at the kth period. It can be expressed as

$$D_{ij}^m(k) = \widetilde{D}_{ij}^m(k) + \sum_{r \neq m} \varphi_{ij}^r \left[D_{ij}^r(k) - S_{ij}^r(k) \right],$$

$$\forall \varphi_{ij}^r \geq 0, r \in M_{ij}. \tag{3}$$

Applying formula (1) into (3), we have

$$D_{ij}^m(k) = \alpha_{ij}^m \left\{ X_{ij}^m(k) \right\}^T + \beta_{ij}^m P_{ij}^m(k)$$
$$+ \sum_{r \neq m} \varphi_{ij}^r \left[D_{ij}^r(k) - S_{ij}^r(k) \right], \tag{4}$$

where φ_{ij}^r represents the spillover effect coefficient between the mth travel market and the rth travel market on OD pair (i, j).

Considering the interaction of the effective supply between different OD pairs in the same travel mode (e.g., a driverless bus passes through node i and node i' to node j; if the effective demand on OD pair (i, j) increases, the effective supply on OD pair (i, j) will also increase, and the effective supply on OD pair (i', j) will decrease), we define the linear combination between the notional demand of the mth travel market and the effective supply of other travel markets as the effective supply S_{ij}^m of the mth travel market on OD pair (i, j) at the kth period. It can be expressed as

$$S_{ij}^m(k) = \widetilde{S}_{ij}^m(k) + \sum_{i' \neq i \text{ or } j' \neq j} \vartheta_{ij-i''j}^m S_{i''j}^m(k), \quad \forall \vartheta_{i''j}^m \leq 0, i', j' \in N. \tag{5}$$

Applying formula (2) into (5), we have

$$S_{ij}^m(k) = \eta_{ij}^m \left\{ Y_{ij}^m(k) \right\}^T + \sum_{i' \neq i \text{ or } j' \neq j} \vartheta_{ij-i''j}^m S_{i''j}^m(k), \tag{6}$$

where $\vartheta_{ij-i''j}^m$ represents the influence coefficient of the mth travel market between OD pair (i, j) and OD pair (i', j') and $\vartheta_{i''j}^m = 0$ indicates that the mth travel markets between OD pair (i, j) and OD pair (i', j') are independent of each other.

2.3. Model Formulation. In the multimodal travel market, if the effective demand D_{ij}^m is less than the effective supply S_{ij}^m, the travel market is in the state of oversupply and the market trading volume $Q_{ij}^m = D_{ij}^m$; if the effective demand D_{ij}^m is more than the effective supply S_{ij}^m, the travel market is in the state of overdemand and the market trading volume $Q_{ij}^m = S_{ij}^m$; if the effective demand D_{ij}^m is equal to the effective supply S_{ij}^m, the travel market is in the state of equilibrium and the market trading volume $Q_{ij}^m = D_{ij}^m = S_{ij}^m$. Clearly, the multimodal travel market trading follows the principle of short side. The trading volume Q_{ij}^m of the mth travel market on OD pair (i, j) at the kth period can be expressed as

$$Q_{ij}^m(k) = \min\left\{ D_{ij}^m(k), S_{ij}^m(k) \right\}. \tag{7}$$

Using formulas (4), (6), and (7), the multimodal travel market model on OD pair (i, j) can be expressed as

$$\begin{cases} D_{ij}^m(k) = \alpha_{ij}^m X_{ij}^m(k) + \beta_{ij}^m P_{ij}^m(k) + \sum_{r \neq m} \varphi_{ij}^r \left[D_{ij}^r(k) - S_{ij}^r(k) \right], \\ S_{ij}^m(k) = \eta_{ij}^m Y_{ij}^m(k) + \sum_{i' \neq i \text{ or } j' \neq j} \vartheta_{ij-i''j}^m S_{i''j}^m(k), \\ Q_{ij}^m(k) = \min\left\{ D_{ij}^m(k), S_{ij}^m(k) \right\}. \end{cases} \tag{8}$$

3. Regulation Mechanism of Disequilibrium Multimodal Travel Market

To regulate the supply-demand relationship of the disequilibrium multimodal travel market, this section attempts to seek the optimal price regulation law based on the multimodal travel market model by using the dynamic programming method.

3.1. Optimal Price Regulation Law. Let the total number of nodes in set N be λ, the total number of travel modes in set M_{ij} is ℓ_{ij}, $D(k) = [D_{12}(k), \dots, D_{1\lambda}(k), D_{21}(k), \dots, D_{\lambda(\lambda-1)}(k)]^T$, $X(k) = [X_{12}(k), \dots, X_{1\lambda}(k), X_{21}(k), \dots, X_{\lambda(\lambda-1)}(k)]^T$, $P(k), = [P_{12}(k), \dots, P_{1\lambda}(k), P_{21}(k), \dots, P_{\lambda(\lambda-1)}(k)]^T$, $S(k) = [S_{12}(k), \dots, S_{1\lambda}(k), S_{21}(k), \dots, S_{\lambda(\lambda-1)}(k)]^T$, and $Y(k) = [Y_{12}(k), \dots, Y_{1\lambda}(k), Y_{21}(k), \dots, Y_{\lambda(\lambda-1)}(k)]^T$. Then, the effective demand and effective supply in the multimodal travel market model (8) can be rewritten as

$$D(k) = \alpha X(k) + \beta P(k) + \varphi[D(k) - S(k)], \tag{9}$$

$$S(k) = \eta Y(k) + \vartheta S(k), \tag{10}$$

where

$$D_{ij}(k) = \left[D_{ij}^1(k), \ldots, D_{ij}^{\ell_{ij}}(k) \right],$$

$$S_{ij}(k) = \left[S_{ij}^1(k), \ldots, S_{ij}^{\ell_{ij}}(k) \right],$$

$$X_{ij}(k) = \left[X_{ij}^1(k), \ldots, X_{ij}^{\ell_{ij}}(k) \right],$$

$$Y_{ij}(k) = \left[Y_{ij}^1(k), \ldots, Y_{ij}^{\ell_{ij}}(k) \right],$$

$$P_{ij}(k) = \left[P_{ij}^1(k), \ldots, P_{ij}^{\ell_{ij}}(k) \right],$$

$$\alpha_{ij} = \mathrm{diag}\left(\alpha_{ij}^1, \ldots, \alpha_{ij}^{\ell_{ij}} \right), \quad \alpha = \mathrm{diag}\left(\alpha_{12}, \ldots, \alpha_{1\lambda}, \alpha_{21}, \ldots, \alpha_{\lambda(\lambda-1)} \right),$$

$$\beta_{ij} = \mathrm{diag}\left(\beta_{ij}^1, \ldots, \beta_{ij}^{\ell_{ij}} \right), \quad \beta = \mathrm{diag}\left(\beta_{12}, \ldots, \beta_{1\lambda}, \beta_{21}, \ldots, \beta_{\lambda(\lambda-1)} \right),$$

$$\eta_{ij} = \mathrm{diag}\left(\eta_{ij}^1, \ldots, \eta_{ij}^{\ell_{ij}} \right), \quad \eta = \mathrm{diag}\left(\eta_{12}, \ldots, \eta_{1\lambda}, \eta_{21}, \ldots, \eta_{\lambda(\lambda-1)} \right),$$

$$\varphi_{ij} = \begin{bmatrix} 0 & \varphi_{ij}^2 & \cdots & \varphi_{ij}^{\ell_{ij}-1} & \varphi_{ij}^{\ell_{ij}} \\ \varphi_{ij}^1 & 0 & \cdots & \varphi_{ij}^{\ell_{ij}-1} & \varphi_{ij}^{\ell_{ij}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \varphi_{ij}^1 & \varphi_{ij}^2 & \cdots & 0 & \varphi_{ij}^{\ell_{ij}} \\ \varphi_{ij}^1 & \varphi_{ij}^2 & \cdots & \varphi_{ij}^{\ell_{ij}-1} & 0 \end{bmatrix}, \tag{11}$$

$$\varphi = \mathrm{diag}\left(\varphi_{12}, \ldots, \varphi_{1\lambda}, \varphi_{21}, \ldots, \varphi_{\lambda(\lambda-1)} \right),$$

$$\vartheta_{ij}^{i''j} = \begin{bmatrix} \vartheta_{ij-i''j}^1 & 0 & \cdots & 0 \\ 0 & \vartheta_{ij-i''j}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \vartheta_{ij-i''j}^{\ell_{i'j}} \end{bmatrix},$$

$$\vartheta = \begin{bmatrix} 0 & \cdots & \vartheta_{12}^{1\lambda} & \vartheta_{12}^{21} & \cdots & \vartheta_{12}^{\lambda(\lambda-1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vartheta_{1\lambda}^{12} & \cdots & 0 & \vartheta_{1\lambda}^{21} & \cdots & \vartheta_{1\lambda}^{\lambda(\lambda-1)} \\ \vartheta_{21}^{12} & \cdots & \vartheta_{21}^{1\lambda} & 0 & \cdots & \vartheta_{21}^{\lambda(\lambda-1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vartheta_{\lambda(\lambda-1)}^{12} & \cdots & \vartheta_{\lambda(\lambda-1)}^{1\lambda} & \vartheta_{\lambda(\lambda-1)}^{21} & \cdots & 0 \end{bmatrix}.$$

Let $E = \mathrm{diag}(1, 1, \ldots, 1)$, $|E - \varphi| \neq 0$ and $|E - \vartheta| \neq 0$. Formula (10) can be abbreviated as

$$S(k) = (E - \vartheta)^{-1} \eta Y(k). \tag{12}$$

Applying formula (12) into (9), we have

$$D(k) = (E - \varphi)^{-1} \left[\alpha X(k) + \beta P(k) - \varphi (E - \vartheta)^{-1} \eta Y(k) \right]. \tag{13}$$

Let the supply-demand deviation $Z(k) = D(k) - S(k)$, we can obtain the following equation according to formulas (12) and (13).

$$Z(k) = (E - \varphi)^{-1} \alpha X(k) - \left[(E - \varphi)^{-1} \varphi (E - \vartheta)^{-1} \eta \\ + (E - \vartheta)^{-1} \eta \right] Y(k) + (E - \varphi)^{-1} \beta P(k). \tag{14}$$

For the convenience of analysis, we suppose that these exogenous variables remain unchanged; that is, $X(k + 1) = X(k)$ and $Y(k + 1) = Y(k)$. Let the regulation amplitude of travel price $\Delta P(k) = P(k + 1) - P(k)$, $U(k) = \Delta P(k)$, $A = E$, and $B = (E - \varphi)^{-1} \beta$; we thus have the following dynamic relational expression from formula (14).

$$Z(k + 1) = AZ(k) + BU(k). \tag{15}$$

To minimize the supply-demand deviation $Z(k)$ and avoid fierce market oscillation caused by excessive travel price regulation, we take the comprehensive minimization between the supply-demand deviation and the regulation amplitude of travel price as the objective function. Let H_0 be a symmetric positive definite matrix, $H = \mathrm{diag}(q_{12}, \ldots, q_{1\lambda}, q_{21}, \ldots, q_{(\lambda\lambda-1)})$, $q_{ij} = \mathrm{diag}(q_{ij}^1, \ldots, q_{ij}^{\ell_{ij}})$, $q_{ij}^1, \ldots, q_{ij}^{\ell_{ij}} > 0$, $R = \mathrm{diag}(\gamma_{12}, \ldots, \gamma_{1\lambda}, \gamma_{21}, \ldots, \gamma_{\lambda(\lambda-1)})$, $\gamma_{ij} = \mathrm{diag}(\gamma_{ij}^1, \ldots, \gamma_{ij}^{\ell_{ij}})$, and $\gamma_{ij}^1, \ldots, \gamma_{ij}^{\ell_{ij}} > 0$. Then, the objective function J can be expressed as

$$\min J = \frac{1}{2} Z^T(L) H_0 Z(L) \\ + \frac{1}{2} \sum_{k=0}^{L-1} \left[Z^T(k) HZ(k) + U^T(k) RU(k) \right]. \tag{16}$$

By using the dynamic programming method (see Appendix) to solve formula (16), we obtain the optimal price regulation law of the disequilibrium multimodal travel market. It can be expressed as

$$P(k + 1) = P(k) - KZ(k), \tag{17}$$

where $K = [B^T GB + R]^{-1} B^T GA$, $H_0 = G$, and G represents the positive definite solution of the following discrete Riccati equation.

$$G = (A - BK)^T G(A - BK) + K^T RK + H \\ = A^T GA - A^T GB(B^T GB + R)^{-1} B^T GA + H. \tag{18}$$

3.2. Stability Analysis

Theorem 1. *For the optimal regulation problem (16), if k is finite, then there exists a unique optimal price regulation law of discrete system (15).*

Proof. We first prove the existence of the optimal price regulation law $U^*(k)$. Because A, B, R, and H are constant matrices, then G is unique according to formula (18); thus, K is also unique. Evidently, there exists at least one optimal price regulation law $U^*(k)$ in discrete system (15).

Second, we prove the uniqueness of the optimal price regulation law $U^*(k)$. Suppose that there exist two different optimal price regulation laws, $U_1^*(k)$ and $U_2^*(k)$, in discrete system (15). From the uniqueness of K, we know the following:

$$U_1^*(k) = -KZ_1(k), \tag{19}$$

$$U_2^*(k) = -KZ_2(k). \tag{20}$$

Applying formulas (19) and (20) into (15), respectively, we have

$$Z_1^*(k + 1) = (A - BK)Z_1(k), \tag{21}$$

$$Z_2^*(k + 1) = (A - BK)Z_2(k). \tag{22}$$

Because Z_1^* and Z_2^* are solutions of the same discrete system with the same initial condition, we know that $Z_1^*(k + 1) = Z_2^*(k + 1)$ from formula (21) and (22); that is, $U_1^*(k) = U_2^*(k)$. Hence, the optimal regulation problem (16) has a unique optimal price regulation law $U^*(k)$. ☐

Theorem 2. *If the optimal regulation problem (16) of discrete system (15) has a unique optimal price regulation law $U^*(k)$, then the optimal price regulation law $U^*(k)$ is asymptotically stable.*

Proof. We first introduce the stability criterion of discrete system (Tian [33]). ☐

3.2.1. Stability Criterion of Discrete System. If there exists a discrete function $V(Z(k))$ for discrete system (15), $V(Z(k))$ satisfies the conditions,

(1) $V(Z(k))|_{Z(k)=0} = 0$ and $V(Z(k))|_{Z(k)\neq0} > 0$,

(2) $\Delta V(Z(k)) = V(Z(k + 1)) - V(Z(k)) < 0$,

then the optimal price regulation law $U^*(k)$ is asymptotically stable.

Constructing a discrete function $V(Z(k))$, it can be written as

$$V(Z(k)) = Z^T(k) GZ(k). \tag{23}$$

According to the positive definite solution G and formula (23), we know that $V(Z(k))$ satisfies condition (1) in the stability criterion of discrete system, and we only need to prove that $\Delta V(Z(k)) < 0$. We know
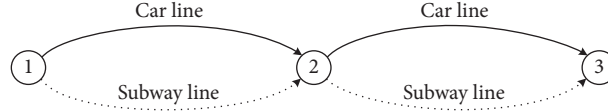
Figure 1: Test road network.

$$\Delta V(Z(k)) = Z^T(k)\left[(A - BK)^T G(A - BK) - G\right]Z(k). \tag{24}$$

According to formula (18), formula (24) can be rewritten as

$$\Delta V(Z(k)) = -Z^T(k)\left[H + K^T RK\right]Z(k). \tag{25}$$

Because H, K, and R are positive definite matrices, we know that $H + K^T RK > 0$; thus, $\Delta V(Z(k)) < 0$. Therefore, the optimal price regulation law $U^*(k)$ is asymptotically stable.

4. Numerical Examples

4.1. Test Road Network and Market Model. To validate the proposed optimal price regulation law, we need to establish the corresponding relationship between travel price in the theoretical model and specific traffic management measures. Hence, we suppose that the test road network includes OD pairs (1,2) and (1,3), driverless car, and driverless subway in Figure 1, in which the travel price of driverless car and driverless subway are regulated by road congestion fee CF and subway fare SF, respectively.

Suppose that the driverless car travel price $P_{ij}^{c_{ij}}(k) = P_{ij}^{c_{ij}}(0) + \mathrm{CF}_{ij}^{c_{ij}}(k)$ and the driverless subway travel price $P_{ij}^{s_{ij}}(k) = \mathrm{SF}_{ij}^{s_{ij}}(k)$; we construct the following multimodal travel market model on OD pairs (1,2) and (1,3).

$$\begin{cases} D_{12}^{c_{12}}(k) = 300 - 2P_{12}^{c_{12}}(k) + 0.2 \times [D_{12}^{s_{12}}(k) - S_{12}^{s_{12}}(k)], \\ S_{12}^{c_{12}}(k) = 0.4 \times 500 + 0.2 \times S_{13}^{c_{13}}(k), \\ Q_{12}^{c_{12}}(k) = \min\{D_{12}^{c_{12}}(k), S_{12}^{c_{12}}(k)\}, \end{cases} \qquad \begin{cases} D_{12}^{s_{12}}(k) = 300 - 25P_{12}^{s_{12}}(k) + 0.2 \times [D_{12}^{c_{12}}(k) - S_{12}^{c_{12}}(k)], \\ S_{12}^{s_{12}}(k) = 0.8 \times 300 + 0.1 \times S_{13}^{s_{13}}(k), \\ Q_{12}^{s_{12}}(k) = \min\{D_{12}^{s_{12}}(k), S_{12}^{s_{12}}(k)\}, \end{cases} \tag{26}$$

$$\begin{cases} D_{13}^{c_{13}}(k) = 350 - 2P_{13}^{c_{13}}(k) + 0.2 \times [D_{13}^{s_{13}}(k) - S_{13}^{s_{13}}(k)], \\ S_{13}^{c_{13}}(k) = 0.4 \times 500 + 0.2 \times S_{12}^{c_{12}}(k), \\ Q_{13}^{c_{13}}(k) = \min\{D_{13}^{c_{13}}(k), S_{13}^{c_{13}}(k)\}, \end{cases} \qquad \begin{cases} D_{13}^{s_{13}}(k) = 360 - 30P_{13}^{s_{13}}(k) + 0.2 \times [D_{13}^{c_{13}}(k) - S_{13}^{c_{13}}(k)], \\ S_{13}^{s_{13}}(k) = 0.8 \times 300 + 0.1 \times S_{12}^{s_{12}}(k), \\ Q_{13}^{s_{13}}(k) = \min\{D_{13}^{s_{13}}(k), S_{13}^{s_{13}}(k)\}, \end{cases} \tag{27}$$

In the effective demand $D_{12}^{c_{12}}$ of driverless car travel market on OD pair (1,2), $X_{12}^{c_{12}} = 300$ represents the total travel demand of driverless car on OD pair (1,2) at the kth hour, $\alpha_{12}^{c_{12}} = 1$ represents the influence of $X_{12}^{c_{12}}$ on $D_{12}^{c_{12}}$, $P_{12}^{c_{12}}$ represents the driverless car travel price on OD pair (1,2) at the kth hour, $\beta_{12}^{c_{12}} = -2$ represents the sensitivity of $P_{12}^{c_{12}}$ on $D_{12}^{c_{12}}$, $D_{12}^{c_{12}} - S_{12}^{s_{12}}$ represents the effective supply-demand difference of driverless subway on OD pair (1,2) at the kth hour, and $\varphi_{12}^{c_{12}} = 0.2$ represents the sensitivity of $D_{12}^{s_{12}} - S_{12}^{s_{12}}$ on $D_{12}^{c_{12}}$.

In the effective supply $S_{12}^{c_{12}}$ of driverless car travel market on OD pair (1,2), $Y_{12}^{c_{12}} = 500$ represents the total supply on OD pair (1,2) at the kth hour, $\eta_{12}^{c_{12}} = 0.4$ represents the ratio of the supply meet driverless car travel demand to the total supply on OD pair (1,2), $S_{13}^{c_{13}}$ represents the effective supply of driverless car travel market on OD pair (1, 3), and $\vartheta_{12-13}^{c_{12}}$ represents the influence of $S_{13}^{c_{13}}$ on $S_{12}^{c_{12}}$.

$D_{12}^{s_{12}}$, $S_{12}^{s_{12}}$, $D_{13}^{c_{13}}$, $S_{13}^{c_{13}}$, $D_{13}^{s_{13}}$, and $S_{13}^{s_{13}}$ are similar to the definitions of the above variables in the models (26) and (27).

4.2. Model Calculation and Result Analysis. Suppose that $H = \mathrm{diag}(1, 1, 1, 1)$, $R = \mathrm{diag}(8, 800, 8, 600)$, the regulation period $k \in [0, 24]$, the regulation step $\Delta k = 1$, the initial

travel price $P_{12}^{c_{12}}(0) = 20$, $P_{12}^{s_{12}}(0) = 2$, $P_{13}^{c_{13}}(0) = 30$, and $P_{13}^{s_{13}}(0) = 4$. The results of calculation are presented in Figure 2 and Table 1.

As shown in Figure 2, the driverless car travel markets on OD pairs (1, 2) and (1, 3) at the beginning are in the state of overdemand, and the driverless subway travel markets are in the state of oversupply. After a certain number of regulation periods, the supply-demand relationship of the driverless car and driverless subway travel markets on OD pairs (1, 2) and (1, 3) reaches the equilibrium state. The results indicate that the proposed optimal price regulation law can significantly reduce the supply-demand deviation of the multimodal travel market and avoid the agitation of travel market caused by excessive price regulation.

The optimal price regulation results of driverless car and driverless subway travel markets on OD pairs (1,2) and (1,3) within 0–9 periods are listed in Table 1. Combining the results of Figure 2 and Table 1 reveals that increasing the road congestion fee and reducing the subway fare will motivate some travelers who plan to travel by driverless car to take the driverless subway or give up their travel plans. Ultimately, it makes the demand of driverless subway travel market increase, makes the demand of driverless car travel market decrease, and
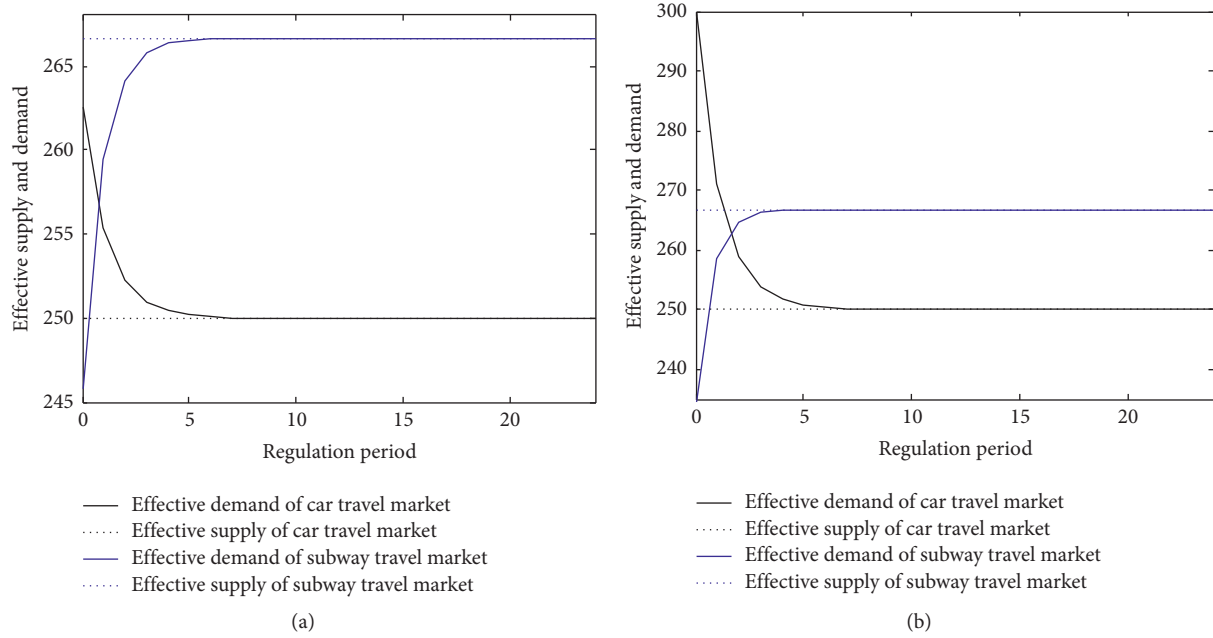
FIGURE 2: Regulation process of effective supply and demand in the multimodal travel market. (a) OD pair (1,2). (b) OD pair (1,3).

TABLE 1: Optimal price regulation results.

Regulation period (h)	OD pair (1, 2)		OD pair (1, 3)	
	Congestion fee	Subway fare	Congestion fee	Subway fare
0	0	2	0	4
1	2.8786	1.5654	11.5146	3.3279
2	4.1000	1.4141	16.3999	3.1640
3	4.6181	1.3615	18.4726	3.1240
4	4.8380	1.3431	19.3520	3.1143
5	4.9313	1.3367	19.7251	3.1119
6	4.9708	1.3345	19.8833	3.1113
7	4.9876	1.3337	19.9505	3.1112
8	4.9948	1.3335	19.9790	3.1111
9	4.9978	1.3334	19.9911	3.1111

finally achieves market equilibrium. The results show that the optimal price regulation law can help travelers to choose a reasonable travel mode to travel.

5. Conclusions

In this study, by taking each type of driverless travel modes on each OD pair as a travel market and taking the driving time, driving cost, and perceived traffic service quality as the travel price, we established a multimodal travel market model. Aiming at achieving the minimum supply-demand deviation as well as the minimum amplitude of price regulation, we proposed an optimal price-quantity law of the disequilibrium multimodal travel market and analyzed the stability of the optimal price regulation law. Through a numerical example, the application of the optimal price regulation law in urban multimodal traffic management was studied. The results indicated that the optimal price regulation law can effectively reduce the supply-demand deviation of the multimodal travel market, avoid the agitation of the travel market caused by excessive price regulation, and achieve market equilibrium.

This study can help travelers choose a reasonable travel mode to travel and promote the development of urban traffic demand management methods. In a future study, we plan to consider the mixed traffic scenario of manual and driverless vehicle, the effects of the travel price on the notional supply in the long period, and the influence of the expected trading volume on traveler' travel mode choice.

Appendix

The processes of solving formula (16) by the dynamic programming method are as follows. Let the $k + 1$ level decision process be a process from $k + 1$ level to the terminal state through $L - k - 1$ level decision-making. Then, the performance index J_{k+1} of the $k + 1$ level decision process can be expressed as

$$J_{k+1} = \frac{1}{2}Z^T(L)H_0Z(L)$$
$$+ \frac{1}{2}\sum_{\tau=k+1}^{L-1}\left[Z^T(\tau)HZ(\tau) + U^T(\tau)RU(\tau)\right]. \quad (28)$$

Hence, the performance index J_k of the k level decision process can be expressed as

$$J_k = J_{k+1} + \frac{1}{2}\left[Z^T(k)HZ(k) + U^T(k)RU(k)\right]. \quad (29)$$

According to Bellman's principle of optimality (Gross [34]), the optimal performance index of formula (29) can be expressed as

$$J_k^* = \min_{U(k)}\left\{J_{k+1}^* + \frac{1}{2}\left[Z^T(k)HZ(k) + U^T(k)RU(k)\right]\right\}. \quad (30)$$

Let G be the symmetric matrix and $J_k^* = 1/2Z^T(k)GZ(k)$. Then, we have

$$J_{k+1}^* = \frac{1}{2}Z^T(k+1)GZ(k+1). \quad (31)$$

By applying formula (15) into (31), we obtain

$$J_{k+1}^* = \frac{1}{2}[AZ(k) + BU(k)]^T G[AZ(k) + BU(k)]. \quad (32)$$

By applying formula (32) into (30), we obtain

$$J_k^* = \min_{U(k)}\left\{Z^T(k)\left[H + A^T GA\right]Z(k) + 2U^T(k)B^T GAZ(k) + U^T(k)\left[B^T GB + R\right]U(k)\right\}. \quad (33)$$

We also know that the regulation variable $U(k)$ has no effect on the current state $Z(k)$ from formula (15). Therefore, we have

$$\frac{\partial J_k^*}{\partial U(k)} = 2B^T GAZ(k) + 2\left[B^T GB + R\right]U(k) = 0. \quad (34)$$

According to formula (34), we obtain the optimal price regulation law of the disequilibrium multimodal travel market, which can be expressed as

$$U^*(k) = -KZ(k). \quad (35)$$

Equivalently,

$$P(k+1) = P(k) - KZ(k). \quad (36)$$

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] M. M. Morando, Q. Y. Tian, L. T. Truong, and H. L. Vu, "Studying the safety impact of autonomous vehicles using simulation-based surrogate safety measures," *Journal of Advanced Transportation*, vol. 2018, p. 11, 2018.

[2] L. Ye and T. Yamamoto, "Impact of dedicated lanes for connected and autonomous vehicle on traffic flow throughput," *Physica A: Statistical Mechanics and Its Applications*, vol. 512, pp. 588–597, 2018.

[3] P. M. Smolnicki and J. Sołtys, "Driverless mobility: the impact on metropolitan spatial structures," *Procedia Engineering*, vol. 161, pp. 2184–2190, 2018.

[4] W. Zhang, S. Guhathakurta, J. Fang, and G. Zhang, "Exploring the impact of shared autonomous vehicles on urban parking demand: an agent-based simulation approach," *Sustainable Cities and Society*, vol. 19, pp. 34–45, 2015.

[5] J. Kamel, R. Vosooghi, J. Puchinger, F. Ksontini, and G. Sirin, "Exploring the impact of user preferences on shared autonomous vehicle modal split: a multi-agent simulation approach," *Transportation Research Procedia*, vol. 37, pp. 115–122, 2019.

[6] Z. Yi, J. Smart, and M. Shirk, "Energy impact evaluation for eco-routing and charging of autonomous electric vehicle fleet: ambient temperature consideration," *Transportation Research Part C: Emerging Technologies*, vol. 89, pp. 344–363, 2018.

[7] A. T. Moreno, A. Michalski, C. Llorca, and R. Moeckel, "Shared autonomous vehicles effect on vehicle-km traveled and average trip duration," *Journal of Advanced Transportation*, vol. 2018, Article ID 8969353, 10 pages, 2018.

[8] Z. Liu and Z. Song, "Strategic planning of dedicated autonomous vehicle lanes and autonomous vehicle/toll lanes in transportation networks," *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 381–403, 2019.

[9] C. Xia, X. Jin, L. Kong, C. Xu, and P. Zeng, "Lane scheduling around crossroads for edge computing based autonomous driving," *Journal of Systems Architecture*, vol. 95, pp. 1–8, 2019.

[10] J. M. L. Domínguez and T. J. M. Sanguino, "Design, modelling, and implementation of a fuzzy controller for an intelligent road signaling system," *Complexity*, vol. 2018, Article ID 1849527, 14 pages, 2018.

[11] N. Jiang, "Optimal signal design for mixed equilibrium networks with autonomous and regular vehicles," *Journal of Advanced Transportation*, vol. 2017, Article ID 5649823, 13 pages, 2017.

[12] M. Tajalli and A. Hajbabaie, "Distributed optimization and coordination algorithms for dynamic speed optimization of connected and autonomous vehicles in urban street networks," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 497–515, 2018.

[13] D. Li, Y. Zhao, P. Ranjitkar, H. Zhao, and Q. Bai, "Hybrid approach for variable speed limit implementation and application to mixed traffic conditions with connected autonomous vehicles," *IET Intelligent Transport Systems*, vol. 12, no. 5, pp. 327–334, 2018.

[14] M. Nourinejad, S. Bahrami, and M. J. Roorda, "Designing parking facilities for autonomous vehicles," *Transportation Research Part B: Methodological*, vol. 109, pp. 110–127, 2018.

[15] R. Estepa, A. Estepa, J. Wideberg, M. Jonasson, and A. Stensson-Trigell, "More effective use of urban space by autonomous double parking," *Journal of Advanced Transportation*, vol. 2017, Article ID 8426946, 10 pages, 2017.

[16] S. Yamashita and K. Takami, "Autonomous, distributed parking lot vacancy management using intervehicle communication," *International Journal of Vehicular Technology*, vol. 2014, Article ID 647487, 9 pages, 2014.

[17] B. Wang, S. A. Ordonez Medina, and P. J. Fourie, "Comparing parking strategies of autonomous transit on demand with varying transport demand," *Procedia Computer Science*, vol. 151, pp. 814–819, 2019.

[18] G. Leich and J. Bischoff, "Should autonomous shared taxis replace buses? a simulation study," *Transportation Research Procedia*, vol. 41, pp. 450–460, 2019.

[19] R. Abe, "Introducing autonomous buses and taxis: quantifying the potential benefits in Japanese transportation systems," *Transportation Research Part A: Policy and Practice*, vol. 126, pp. 94–113, 2019.

[20] Z. Cao and A. Ceder, "Autonomous shuttle bus service timetabling and vehicle scheduling using skip-stop tactic," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 370–395, 2019.

[21] L. Tong, L. Zhou, J. Liu, and X. Zhou, "Customized bus service design for jointly optimizing passenger-to-vehicle assignment and vehicle routing," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 451–475, 2017.

[22] L. Ye and T. Yamamoto, "Evaluating the impact of connected and autonomous vehicles on traffic safety," *Physica A: Statistical Mechanics and Its Applications*, vol. 526, Article ID 121009, 12 pages, 2019.

[23] N. Virdi, H. Grzybowska, S. T. Waller, and V. Dixit, "A safety assessment of mixed fleets with connected and autonomous vehicles using the surrogate safety assessment module," *Accident Analysis & Prevention*, vol. 131, pp. 95–111, 2019.

[24] H. Prakken, "On the problem of making autonomous vehicles conform to traffic law," *Artificial Intelligence and Law*, vol. 25, no. 3, pp. 341–363, 2017.

[25] C. Bartolini, T. Tettamanti, and I. Varga, "Critical features of autonomous road transport from the perspective of technological regulation and law," *Transportation Research Procedia*, vol. 27, pp. pp791–798, 2017.

[26] P. Wagner, *Traffic Control and Traffic Management in a Transportation System with Autonomous Vehicles*, Springer, Berlin, Germany, 2016.

[27] T. Tettamanti, A. Mohammadi, H. Asadi, and I. Varga, "A two-level urban traffic control for autonomous vehicles to improve network-wide performance," *Transportation Research Procedia*, vol. 27, pp. 913–920, 2017.

[28] R. Kitamura, S. Fujii, and E. I. Pas, "Time-use data, analysis and modeling: toward the next generation of transportation planning methodologies," *Transport Policy*, vol. 4, no. 4, pp. 225–235, 1997.

[29] R. Brideges, *Our Driverless Future: Heaven or Hell?*, Understanding Disruption, Inc., Great Britain, 2018.

[30] Y. Ye and H. Wang, "Optimal design of transportation networks with automated vehicle links and congestion pricing," *Journal of Advanced Transportation*, vol. 2018, Article ID 3435720, 12 pages, 2018.

[31] X. Zhang, W. Liu, S. T. Waller, and Y. Yin, "Modelling and managing the integrated morning-evening commuting and parking patterns under the fully autonomous vehicle environment," *Transportation Research Part B: Methodological*, vol. 128, pp. 380–407, 2019.

[32] F. X. Browne, "Sluggish quantity adjustment in a non-clearing market-a disequilibrium econometric application to the loan market," *Journal of Applied Econometrics*, vol. 2, no. 4, pp. 335–349, 1987.

[33] X. G. Tian, "On going further into the extended stability theorem for discrete system," *Journal of Tianjin University*, vol. 28, no. 3, pp. 348–353, 1995.

[34] E. Gross, "On the Bellman's principle of optimality," *Physica A: Statistical Mechanics and Its Applications*, vol. 462, pp. 217–221, 2016.

WILEY | Hindawi

Research Article

Impact of Carriage Crowding Level on Bus Dwell Time: Modelling and Analysis

Yiming Bie,[1] **Yunhao Wang,**[2] **and Le Zhang** ID [3]

[1]*School of Transportation, Jilin University, Changchun 130022, China*
[2]*School of Mathematics and Statistics, Northeast Normal University, Changchun 130024, China*
[3]*Department of Architecture and Civil Engineering, Chalmers University of Technology, Göteborg 412 96, Sweden*

Correspondence should be addressed to Le Zhang; lezh@chalmers.se

This paper develops two types of estimation models to quantify the impacts of carriage crowding level on bus dwell time. The first model (model I) takes the crowding level and the number of alighting and boarding passengers into consideration and estimates the alighting time and boarding time, respectively. The second model (model II) adopts almost the same regression method, except that the impact of crowding on dwell time is neglected. The analysis was conducted along two major bus routes in Harbin, China, by collecting 640 groups of dwell times under crowded condition manually. Compared with model II, the mean absolute error (MAE) of model I is reduced by 137.51%, which indicates that the accuracy of bus dwell time estimation could be highly improved by introducing carriage crowding level into the model. Meanwhile, the MAE of model I is about 3.9 seconds, which is acceptable in travel time estimation and bus schedule.

1. Introduction

1.1. Background. Bus dwell time is defined as the duration of transit vehicle stopped for serving passengers. It includes the total passenger boarding and alighting time and the time needed to open and close doors. In recent years, the use of advanced traffic detection technique, such as the vehicle navigation system based on GPS and automatic passenger counter, offers a far more convenient and efficient data source, which makes it feasible to conduct the transit travel pattern and reliability analysis reasonably [1–5].

Previous studies have shown that dwell time is an important determinant of transit system performance and service quality in many forms of urban public transportation [6–8]. From the perspective of static bus schedule, bus dwell time at stops is a major component of vehicle travel time, while bus travel time plays an important role in determining the departure frequency and route design for public transit planners and operators. The proportion of bus dwell time in total running time can consume up to 26% for some high-frequency, high-ridership bus transit routes, especially in

some high-density areas [9]. Thus, estimating the bus dwell time accurately contributes a lot to planning bus schedule reasonably and efficiently, including the departure headway and the fleet size required to provide service. In terms of dynamic bus schedule, estimating the bus dwell time in advance is of great importance to predict the operating state of vehicles and then determine the punctuality rate or bus headway deviation, avoiding the instability of bus systems or bunching phenomenon. It is also important in determining transit assignment models [10] and reliability analysis of the transit network. Experience also shows that bus arriving irregularly at stops is largely due to the inaccurate estimation of bus dwell time. In addition, while serving passengers at a bus stop, the interaction among buses and passengers may constraint the discharge flows of buses and degrade the bus systems' service quality overall [11, 12]. Consequently, the bus dwell time estimation is essential for improving the service quality as well as reliability of the public transit system [13–16].

There is no doubt that bus dwell time is affected by various factors, including the most significant contributing

factor, that is, the number of boarding or/and alighting passengers, and some secondary contributing factors, for example, the type of bus stops (curbside bus stops or bus bays); payment methods (cash, magnetic stripe tickets, or electronic smart media card); platform height as well as door width. Meanwhile, crowding is also a nonnegligible factor in estimating bus dwell time and longer marginal boarding or/and alighting time is associated with the increase in carriage crowding. It has been well recognized that when buses dwell to load and unload passengers, boarding and alighting occur more slowly when standees are present, especially under door crowding conditions. The standees in the bus or/and around the door impede the boarding (alighting) passengers moving towards the carriage (door) from the bus door (carriage). In other words, the amount of space available decides the circulation speed of passengers within the vehicle. However, higher crowding level does not always result in longer dwell time. For example, when a bus stop is located around a supermarket and also near the terminal station, then fewer passengers are waiting at the stop and most passengers would get off, which accelerates the alighting process accordingly. Consequently, the bus dwells at this stop shorter even if it is highly crowded.

As far as the author knows, the systematical analysis of crowding level on bus dwell time is still missing. This paper aims at studying how carriage crowding level affects dwell time and develops a novel estimation model, taking into account the effect of crowding level for bus dwell time estimation.

1.2. Literature Review.

In the past thirty years, much attention has been paid on the subject of bus dwell time estimation model. According to the influence factors, estimation models proposed by previous studies can be categorized into three classes as follows.

1.2.1. Models considering Alighting and Boarding Passenger Volume.

The number of boarding and alighting passengers is regarded as the major determinant of bus dwelling. Under given average boarding or alighting speed, the larger the number of passengers disembarking from the bus or waiting at the stop is, the longer buses will dwell [17, 18]. Thus, multiple regression models were developed to analyze the influence of alighting and boarding passenger volume based on data collected from surveys [19–21].

For example, Rajbhandari et al. [9] obtained the data from automatic passenger counter, and the impact of boarding and alighting passengers on dwell time was investigated. Li et al. [22] used data collected from Florida's Broward County Transit system to develop a binary door choice model predicting the proportion of alighting passengers who used the front or rear door to disembark from the bus; then a dwell time estimation model was developed. Li et al. [23] proposed two dwell time models for the BRT station. The first model was a linear model while the second was nonlinear. They introduced the conflict between passengers boarding and alighting into the models. González et al. [24] proposed a bus dwell time model obtained by

means of a robust statistical evaluation of boarding passenger data at stops. In the model, dwell time was not increased at a fixed rate of time per passenger. Rashidi and Ranjitkar [25] assessed four different time series based methods (random walk, exponential smoothing, moving average, and autoregressive integrated moving average) to model and estimate bus dwell time based on AVL data collected from Auckland. Rashidi and Ranjitkar [26] proposed a gene expression programming-based approach to model and estimate bus dwell time. The proposed model was calibrated and validated using the data collected from 22 bus stops.

1.2.2. Models considering Secondary Factors.

Although dwell time is highly correlated with the number of boarding and alighting passengers, there are also some secondary contributing factors to the bus dwell time, such as crowding, fare type, platform, and bus design [27, 28]. These secondary factors may strongly influence the effectiveness of different strategies used to improve service. Guenthner and Hamat [29] investigated nine different fare types and payment methods and found that they did not affect the dwell time significantly. Levine and Torng [30] evaluated the impacts of low-floor bus design on reducing bus dwell time. Milkovits [31] used the data from the automatic passenger counting, automatic fare counting, and automatic vehicle location systems installed on Chicago Transit Authority buses, then developed a dwell time estimation model, and analyzed the impact of the secondary factors. Currie et al. [32] analyzed the impacts of platform design on streetcar dwell time based on the data collected from Melbourne, Australia, and Toronto, Canada. Jaiswal et al. [33] considered effects of passengers walking on a relatively longer BRT station platform. Results showed that the long BRT platform may lead to the bus experiencing a higher dwell time. Fernández et al. [34] evaluated the impacts of platform height, door width, and fare collection method on bus dwell time based on filed observed data. Tirachini [35] developed multiple regression models to explain the relationship between observed variation in dwell times and different payment methods, the existence of steps at doors, the age of passengers, and the possible friction between users boarding, alighting, and standing.

1.2.3. Models considering Bus Stop Types.

All studies mentioned above assumed that bus dwell time is highly related to passenger alighting and boarding volume or velocity, while the third category indicated that it was also affected by bus stop design. Meng and Qu [36] pointed out that the bus dwell time at a bus bay was affected by interactions among buses, arrival passengers, and traffic on the shoulder lane. Bus dwell time at a bus bay possessed a high degree of uncertainty originating from the merging behavior of bus to the vehicles in the shoulder lane. A novel probabilistic methodology was developed to estimate the bus dwell time, including a standard regenerative stochastic process to model the interactions among buses, arrival passengers, and traffic on the shoulder lane.

Besides the three types of studies discussed above, Katz and Garrow [37] investigated how bus design factors influence door crowding and quantified how door crowding is related to operational performance and passenger safety. Results showed that door crowding was affected by multiple bus design factors, including door placement, aisle length, presence of a front seating area, and service type. Increases in door crowding were associated with longer marginal boarding times. Fletcher and El-Geneidy [38] determined the influence of crowding and fare payment on dwell time through manual data collection. Multiple regression dwell time models were performed by using a traditional model and a new expanded model. The expanded model showed that crowding significantly increased dwell time after approximately 60% of bus capacity was surpassed. This was the only study considering the effect of bus crowding on dwell time; however, the quantitative method of crowding level was not given in this paper.

1.3. Research Contribution. In this paper, we propose a quantitative model of crowding level and develop an estimation model of bus dwell time, stressing the influence of carriage crowding level. It will be significant in improving the accuracy of bus dwell time estimation. This research can help public transit planners develop better policies associated with crowding.

This paper is organized as follows. Section 2 presents the data collection method; the quantitative model of crowding level and estimation model of bus dwell time and comparison with the conventional estimation model are proposed in Section 3; and conclusions are presented in Section 4.

2. Data Collection

The data used in this research come from two high-frequency bus lines of Harbin, Route 8 and Route 63, by manual data collection, as shown in Figure 1. Both routes pass through the central area of this city. Route 8 is 13.5 km with 25 bus stops, connecting Harbin Children's Hospital with Minjiang Community, a popular residential suburb. Another 10.1 km bus line, Route 63, has 21 bus stops and terminates at Dajiang Community, which is another high-density residential area. All the passengers are required to board from the front door and alight by the rear door. Thus the dwell time of a bus is determined by the maximum value between passenger boarding time and alighting time. To best capture the effects of crowding, data used in this paper were collected during the morning (7:00 am–9:30 am) and afternoon (4:30 pm–7:00 pm) peak hour periods, 19–23 December, 2016. Meanwhile, if bus stop is close to downstream intersection, then bus dwell time will also be affected by bus queuing state at intersection. To eliminate the disturbance of traffic signals, we examined the distribution of all the stops along Route 8 and Route 63. As the distance from each stop to the corresponding downstream intersection is all no less than 100 meters, the influence of traffic signals can be ignored in this study. In addition, control

strategies at stops like bus holding to keep regular headway have not been applied in Harbin, so the effect of holding on bus dwelling does not exist in this study.

Two investigators are arranged for each bus, recording boarding and alighting data individually. The recorder near the front door makes notes of the moment of opening the door, the number of boarding passengers, boarding time, and the moment of closing the door. The other one records the corresponding alighting information of the rear door. The tickets for both routes are one yuan and passengers are required to pay when boarding. As 95% of passengers choose to pay by cash and no change are offered by each bus along these two routes, the effect of payment methods on bus dwell time can be negligible in this paper.

In general, we regard the boarding or alighting time as the difference value between closing and opening moment of the front or rear door. However, it is not always accurate. Previous studies have shown that bus dwell time is also associated with the type of bus stops as well as road traffic status. For example, after the bus loads or unloads all the passengers at a bus bay, it would leave immediately. However, before joining into the traffic, the bus has to spend some time in finding an acceptable gap between consecutive vehicles on the left lane, resulting in longer dwell time. In this case, the real moment for closing door is set as the sum of the moment of the last passenger boarding or alighting the vehicle plus the average duration for closing door, where the average duration for closing a door can be calculated by previous data in general cases; that is, nonpassenger related delays are not included in bus dwell time in our study.

The traffic condition in Harbin is always affected by low temperature snow and ice weather. To eliminate this effect, before the recorders boarded the bus, the weather, temperature, and date were all documented. All the collected data were analyzed preliminarily and abnormal data were removed before estimation, that is, the engine failure caused by severe weather and wheelchair ramp event. We collected 828 groups of bus dwell time at stops. As during peak hour period, the headway of these two lines is short, about 4 minutes, and passenger arrival rate differs from stops. The bus crowding level also varies from each other. We also need to remove the data where no standees are occurred. Finally, we get 640 groups of valid data in total used for Section 3.

3. Model Development and Evaluation

In this section, we develop two estimation models. The first model takes the interaction of bus crowding and number of alighting and boarding passengers into consideration and estimates the alighting time and boarding time in Section 3.1. The second model is indeed a traditional one, only considering the number of boarding and alighting passengers as shown in Section 3.2. Section 3.3 compares the performance of these two models.

3.1. Development of Model I

3.1.1. Crowding Level. Generally, a vehicle is in a crowded state when people on the vehicle impede the individuals
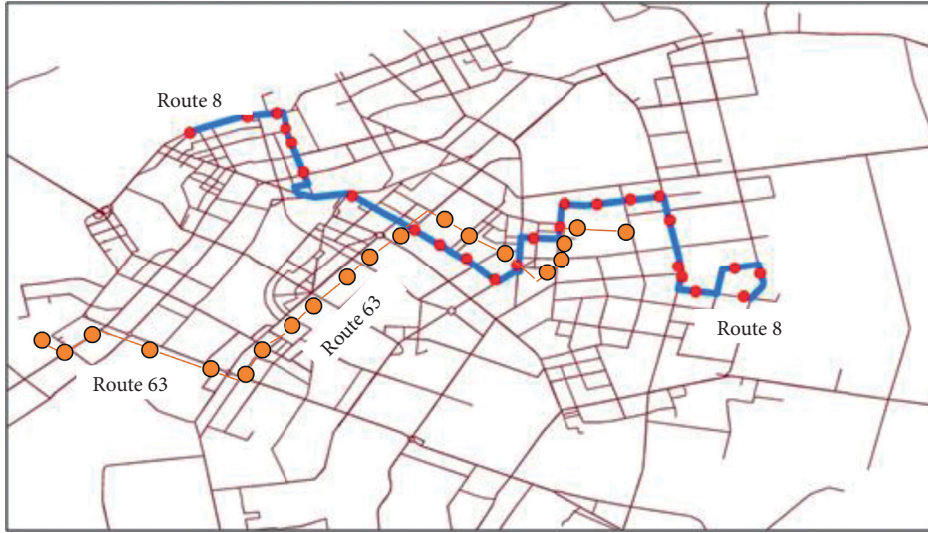
FIGURE 1: Spatial distribution of bus stop on Route 8 and Route 63.

boarding and alighting. Buses on different routes may have different characteristics, such as the space and capacity (seated and standing) for passengers. To truly measure the effect of crowding on bus stopping duration, a more refined formulaic definition is needed. We define crowding level C quantitatively as follows:

$$C = \frac{N_1}{S_1 \times \lambda}, \tag{1}$$

where N_1 is the number of standees in the vehicle and calculated by subtracting the number of seats from the number of passengers on board; S_1 denotes the available space for standees which is determined by the specialty of buses; and λ is the maximum number of standees in unit space of the bus, passengers/m². According to our investigation in Harbin city, unit space of buses can accommodate 7 standees at maximum; that is, $\lambda = 7$ passengers/m².

Next, we will estimate alighting time and boarding time in Section 3.1.2 and the longer one plays a decisive role in bus dwelling duration in Section 3.1.3.

3.1.2. Boarding Time and Alighting Time Estimation Models

(1) Boarding Time Estimation Model. Our objective is to estimate the dynamic relationship of passenger boarding time with bus crowding level and the number of boarding passengers. Firstly, 6 types of regression models are developed to fit the collected 640 data sets and their performance measured by the adjusted coefficient of determination (Adj R^2) is shown in Table 1, where independent variables include the number of boarding passengers (X_1) and bus crowding level (C); boarding time (Y_1) is the dependent variable.

Here, we select the Adj R^2 as evaluative criteria, instead of coefficient of determination, R^2. This is because R^2 may increase spuriously when extra explanatory variables are added. Consequently, to improve the accuracy of estimation, some irrelevant variables may be introduced into the model,

resulting in over fitting. The Adj R^2 can avoid this phenomenon, as the number of parameters (K) is also taken into consideration as shown in (2).

The Adj R^2 indicates the proportion of the variation explained by the estimated regression model, defined as follows:

$$\text{Adj } R^2 = 1 - \frac{\text{SSE}/(n-K)}{\text{SST}/(n-1)} = 1 - \frac{n-1}{n-K}\left(1 - R^2\right), \tag{2}$$

where n is the size of random sample; SSE and SST are the sum of squared errors and the sum of squares of statistical sample respectively; K indicates the number of variables.

According to Table 1, the fitting result of double-logarithmic regression model is the best, as the model explains 74.7% of the variation, which is higher than other models. The exact double-logarithmic regression model is presented in

$$\ln Y_{1i} = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln C_i + e_i, \quad 0 < C_i \le 1, \tag{3}$$

where i is the index of observation; β_0, β_1, and β_2 are parameters to be estimated; e_i denotes disturbance of observation i.

Then we use ordinary least squares method (OLS) to estimate the parameters of (3) and the estimation results are shown in Table 2.

Traditional OLS method requires that there is the constant variance in the errors over all values of the explanatory variables (which is also called homoscedasticity). Once this condition is not satisfied, the validity of parameter estimation cannot be guaranteed. If heteroscedasticity occurs, this estimation is invalid even that the regression coefficient of (3) is evident at the 5% significant level. Then, we need to examine whether the variance of errors is constant or not by some statistical tests, for example, White test and Breusch-Pagan test. Here, we use White test to judge homoscedasticity or not.

To test for a constant variance, an auxiliary regression analysis is conducted, which regresses the squared residuals

TABLE 1: Boarding time models and performances.

Function type	Adj R^2
Linear	0.666
Quadratic	0.670
Cubic	0.670
Exponential	0.605
Semilogarithmic	0.644
Double logarithmic	0.747

TABLE 2: Parameter estimations of boarding time model based on OLS.

Parameters	β_0	β_1	β_2
Estimated value	0.958	0.929	0.086
t-statistic	8.073	18.489	2.307
p-value	<0.001	<0.001	0.023

F-statistic = 194.334 (p-value <0.001); R^2 = 0.751; Adj R^2 = 0.747.

from the original regression model onto a set of regressors as shown in (4). The basic logic of the White test is to substitute the squared residuals for the variance of the error term at each observation. The squared residuals from the original model serve as a proxy for the variance of the error term at each observation. If the error term in the original model has a constant variance in fact, then the coefficients in the auxiliary regression should be statistically indistinguishable from zero and F-statistic should be small. Otherwise, evident F-statistic counts against the hypothesis of homoscedasticity:

$$\widehat{e}_i^2 = \gamma_0 + \gamma_1 \left(\ln X_{1i}\right)^2 + \gamma_2 \left(\ln C_i\right)^2, \quad 0 < C_i \leq 1, \quad (4)$$

where $\widehat{e}_i$ is the estimated value of e_i in (3), that is, residual.

White test is applied to (3) based on the auxiliary regression in (4); we get F-statistic which is 4.461 and p-value = 0.013. Thus at the 5% significant level, homoscedasticity is not satisfied and we need to modify (3).

Generally, the method of weighted least squares (WLS) can be used when the OLS assumption of constant variance in the errors is violated. It works by incorporating extra nonnegative weights associated with each data point into the fitting criterion. The value of weight indicates the precision of information contained in the associated observation. Usually, an observation with small error variance has a large weight as it contains relatively more information than an observation with large error variance. The objective function of ordinary least squares method is to minimize $\sum_{i=1}^{n} \left(Y_{1i} - \widehat{Y}_{1i}\right)^2$, where $\widehat{Y}_{1i}$ is the fitted value of Y_{1i} and n denotes the size of random sample. Accordingly, WLS method aims at minimizing $\sum_{i=1}^{n} w_i \left(Y_{1i} - \widehat{Y}_{1i}\right)^2$ and w_i is the weight of ith observation. Without loss of generality, we define $1/|e_i|$ as the weight, that is, $w_i = 1/|e_i|$, and get the parameter estimations under WLS method as shown in Table 3.

For the modified model, the White test shows that parameter estimations are valid and all the regression coefficients are evident at 5% significant level. Meanwhile, the adjusted coefficient of determination also increases to 0.938, indicating that the accuracy of regression by WLS method is

highly improved compared with that estimated by OLS method. Table 3 shows that 1% increase in the number of boarding passengers will lead to 0.926% increase in boarding time. In the meantime, the crowding level also reveals a positive relationship with boarding time in which a 1% increase in crowding level improves boarding time by 0.085%. As a secondary contributing factor, the effect of crowding level reaches up to almost 10% of the influence caused by the most contributing factor, that is, the number of boarding passengers in bus boarding time, which also specifies the significant importance of carriage crowding level on bus dwell time estimation.

Figure 2(a) shows the relationship between the number of boarding passengers (X_1) and the boarding time (Y_1) under given crowding level (C), where different curves represent different crowding levels. The figure shows that boarding time increases as the number of boarding passengers grows, which is as expected. In the meantime, when the number of boarding passengers is fixed, the boarding time also increases as the vehicle becomes more crowded, which is consistent with the previous parameter analysis. The figure also shows that larger crowding level indicates the higher growth rate of bus dwell time. To examine how the crowding level affects the bus boarding time, we plot the relationship between the number of boarding passengers and bus crowding level in Figure 2(b). When the number of boarding passengers is small, boarding time is indeed insensitive to crowding level. However, when X_1 is up to 25 persons, the boarding time is 42.5 seconds under 0.1 crowding level and reaches up to 51.3 seconds under 0.9 crowding level, increased by 20.7%.

(2) Alighting Time Estimation Model. We will use the similar method presented above to estimate the alighting time (Y_2) with independent variables: the number of alighting passengers (X_2) and bus crowding level (C). The double-logarithmic model also performs best among these six types of models with the Adj R^2 equal to 0.678.

The exact alighting time model adopting double-logarithmic approach is presented in

$$\ln Y_{2i} = \beta_0' + \beta_1' \ln X_{2i} + \beta_2' \ln C_i + e_i, \quad 0 < C_i \leq 1. \quad (5)$$

The parameter estimation of (5) based OLS method shown in Table 4 is valid. The regression coefficients are significant at 5% level. It also reveals that 1% increase in the number of alighting passengers will lead to 0.848% increase in alighting time. Meanwhile, a 1% increase in crowding level improves alighting time by 0.092%, similar parameter analysis results to boarding process:

$$\widehat{e}_i^2 = \gamma_0 + \gamma_1 \left(\ln X_{2i}\right)^2 + \gamma_2 \left(\ln C_i\right)^2, \quad 0 < C_i \leq 1. \quad (6)$$

Comparing the estimation models for boarding time and alighting time, the former presents heteroscedasticity in errors while homoscedasticity is observed in the latter. It is reasonable as the boarding process is disturbed by more various factors, such as the aging of boarding passengers, fare payment method, and the activities of boarding passengers. All the disturbing factors will accumulate and

TABLE 3: Parameter estimations of boarding time model based on WLS.

Parameters	β_0	β_1	β_2
Estimated value	0.965	0.926	0.085
t-statistic	148.700	298.038	41.784
p-value	<0.001	<0.001	<0.001

F-statistic = 47686.58 (p-value <0.001); R^2 = 0.942; Adj R^2 = 0.938.
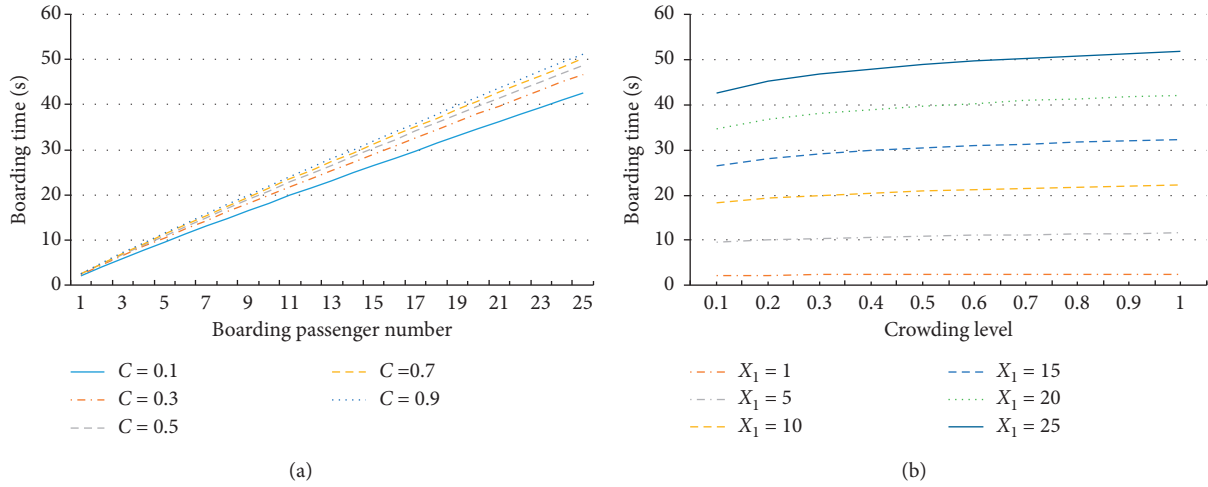

FIGURE 2: The effect of crowding level and boarding passenger volume on boarding time. (a) Relation between boarding passenger number and boarding time under given crowding level. (b) Relation between crowding level and boarding time under given boarding passengers.

TABLE 4: Parameter estimations of alighting time model based on OLS.

Parameters	β_0'	β_1'	β_2'
Estimated value	0.635	0.848	0.092
t-statistic	4.996	16.230	2.503
p-value	<0.001	<0.001	0.014

F-statistic = 150.289 (p-value <0.001); R^2 = 0.682; Adj R^2 = 0.678.

become more distinctive as buses load more passengers. However, alighting process is free of such disturbing factors largely.

Figure 3(a) clearly illustrates the effect of alighting passenger number on alighting time under different levels of crowding level. The curves reveal that the alighting time increases as more passengers get off. However, when the number of alighting passengers is larger, higher crowding level indicates longer alighting time but with smaller increasing range. For example, when 24 passengers alight, the alighting time when $C = 0.3$ is much longer than that under 0.1 crowding level, but the difference of alighting time is much smaller between 0.9 crowding level and 0.7 crowding level even that the gaps between the crowding level keep the same. This is because when more passengers alight, a lot of standees will also get off, which therefore does not cause a barrier to alighting passengers.

The relationship between the alighting time and bus crowding level is presented in Figure 3(b). Similar to Figure 2(b), when the number of alighting passengers is small, alighting time is also insensitive to carriage crowding, as compared with the alighting passenger number; the effect of crowding is only secondary contributing to alighting time. However, when the number of alighting passengers is up to 25 persons, the alighting time is 23.4 seconds under 0.1 crowding level and increases to 28.6 seconds under 0.9 crowding level, increased by 22.4%. Further analysis reveals that average alighting time is much shorter than average boarding time under the same crowding level, which is also expected as paying fare is also time-consuming.

3.1.3. Dwell Time Estimation Model. Obviously, bus dwell time is determined by the maximum value between passenger boarding time and alighting time. For each stop, we choose the maximum one between the estimated boarding time and alighting time to fit bus dwell time (Y) linearly as shown in (7) and the parameter estimations are summarized in Table 5:

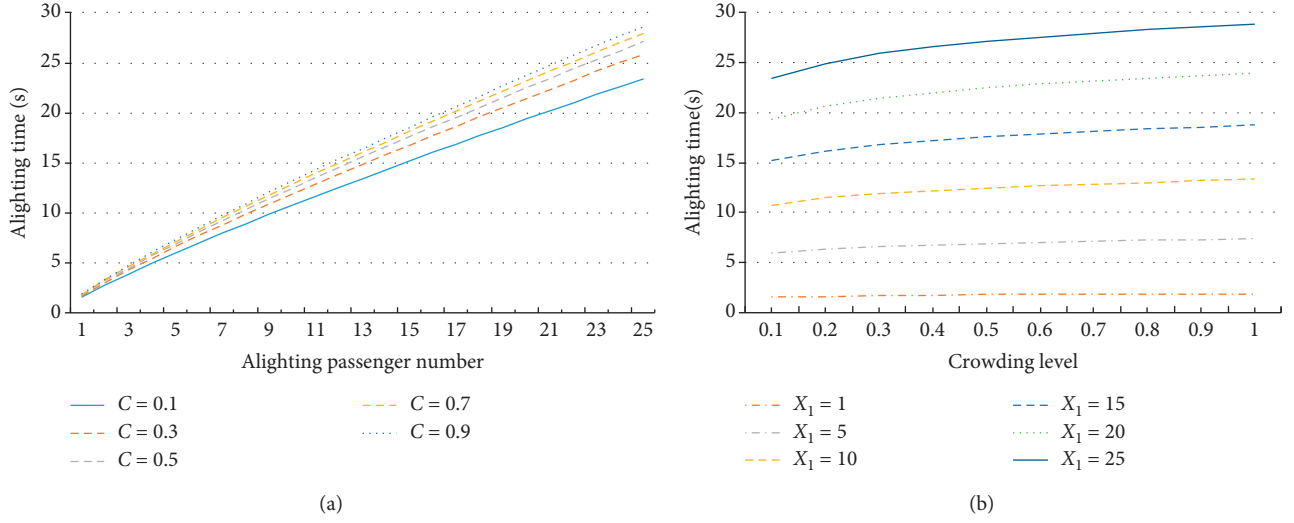$$Y_i = \beta_0'' + \beta_1'' \max\{Y_{1i}, Y_{2i}\} + e_i. \tag{7}$$

FIGURE 3: The effect of crowding level and alighting passenger volume on alighting time. (a) Relation between alighting passenger number and alighting time under given crowding level. (b) Relation between crowding level and alighting time under given alighting passengers.

TABLE 5: Parameter estimations of bus dwell time model based on OLS.

Parameters	β_0''	β_1''
Estimated value	6.819	0.968
t-statistic	8.092	20.320
p-value	<0.001	<0.001

F-statistic = 412.886 (p-value <0.001); R^2 = 0.728; Adj R^2 = 0.727.

TABLE 6: Parameter estimations of bus dwell time model based on WLS.

Parameters	β_0''	β_1''
Estimated value	6.936	0.947
t-statistic	65.799	131.663
p-value	<0.001	<0.001

F-statistic = 17335.21 (p-value <0.001); R^2 = 0.954; Adj R^2 = 0.950.

Similarly, to guarantee the homoscedasticity property, the White test is conducted and an auxiliary regression analysis is developed as shown in

$$\widehat{e}_i^2 = \gamma_0 + \gamma_1 \max\{Y_{1i}, Y_{2i}\}. \tag{8}$$

We get F-statistic equal to 7.696 and p-value is 0.006, implying that the homoscedasticity property does not hold at the 5% significant level. It also shows that bus dwell time is significantly influenced by the interference factors along boarding and alighting process.

We set $1/|e_i|$ as weight as Section 3.1.1 and get the parameters estimations by WLS method listed in Table 6.

The coefficient is evident at the 5% significant level and the Adj R^2 is high to 0.954, implying that the novel expanded model can explain 95.4% of the variation by using a sample size of 640 dwells. Model (7) also shows that bus dwell time consists of two parts: one is a constant value, 6.936 seconds, including the time used for opening and closing doors, as well as the preparation time for passengers to alight and board; the other part is the time used for loading and unloading passengers.

To further confirm whether the constant part is 6.936 seconds or not, we collected 640 data recording the duration including the passenger preparation time and door opening and closing time, denoted as Z_i. The mean of this sample is $\overline{Z}$ = 7.443 and standard deviation is σ = 5.324. It is desired to test the simple hypothesis H_0: μ = 6.936. We propose an

estimator as $t = \overline{Z} - \mu/(\sigma/\sqrt{\mu})$ and we get $t = 1.118 < t_{0.025}(n-1) = 1.98$. Thus, at the 5% significant level, the hypothesis cannot be rejected, and it is reasonable to get the result that the total amount of time required for bus doors opening and closing as well as passenger preparation is about 6.936 seconds, which can also be approximated as 7 seconds.

3.2. Development of Model II.

To clearly show the effect of crowding in bus dwell time, model II is established in almost the same way as described in Section 3.1, except for the crowding level which is not considered in this model; that is, model II describes the relationship between the number of boarding and alighting passengers and bus dwell time.

According to the method described in Section 3.1.1, we get the estimation model of boarding time and alighting time as shown in (9) and (10), respectively:

$$\ln Y_{1i} = 0.736 + 0.973 \ln X_{1i}, \tag{9}$$

$$\ln Y_{2i} = 0.416 + 0.875 \ln X_{2i}. \tag{10}$$

Thus, the bus dwell time estimation model based on model (9) and model (10) is listed as follows:

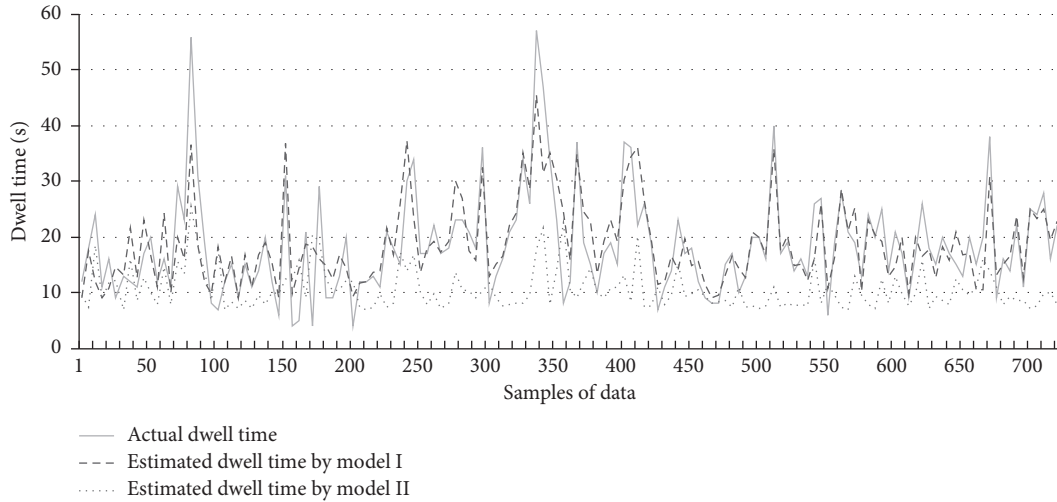$$Y_i = 6.936 + 0.968 \max\{Y_{1i}, Y_{2i}\}. \tag{11}$$

FIGURE 4: Comparison between real and estimated dwell times based on model I and model II.

3.3. Comparisons of the Two Models. To improve the accuracy of bus dwell time estimation model and analyze the effect of bus crowding level, a comparison is conducted between the extended and the traditional model in this section.

(1) Comparisons of Estimation Errors. Figure 4 unveils the predicted dwell time estimated by model I (described by red curve) and model II (described by blue curve) compared with the actual recorded bus dwell time denoted by the black curve. Among the sample with 640 independent data, the maximum bus dwell time is 57 seconds, while the minimum is only 4 seconds. Figure 4 clearly shows that model I fits the real dwell time better. Next, we compare the absolute error of these two models in Figure 5, which indicates the comparisons of predicted value versus observed value. In Figure 5, red line is the absolute error of model I and the blue curve represents that of model II.

Obviously, we can directly see that model I performs much better than model II, as the absolute error of model I is much smaller than that of model II. To measure the overall performance more accurately, we study the mean absolute error (MAE) of these two models as defined by

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left| Y_i - \widehat{Y}_i \right|, \tag{12}$$

where n is the size of random sample; Y_i indicates the actual time; and $\widehat{Y}_i$ is the estimated value of Y_i. Here, Y_i may represent passenger boarding time, alighting time, or bus dwell time.

The MAE of expanded models and traditional models are summarized in Table 7.

Table 7 shows that MAE of dwell time between model II and actual bus dwell time is up to 9.3 seconds, while the MAE caused by model I is only 3.9 seconds averagely, reduced by 137.51%. It unveils that bus crowding level is a nonnegligible factor in bus dwell time estimation and taking the carriage crowding into consideration can highly improve the accuracy of the estimation model for bus dwell time. Meanwhile, we also notice that the MAE of model I (about 3.9 seconds on average) is acceptable in both static and dynamic bus schedule.

In Table 8, we display 6 groups of the collected survey data at stops, including the number of boarding and alighting passengers, carriage crowding level, and actual bus dwell time. It reveals that the carriage crowding level plays a vital role in bus dwell time when the number of boarding and alighting passengers at different stops is the same, which is consistent with finding derived from Table 7.

(2) Analysis on Absolute Errors. In model I, the absolute errors of 391 data sets are less than 3 seconds, taking 61.9% of collected samples. The data with absolute error larger than 5 seconds occupy a fairly small proportion, 22.5%. Among 640 groups of collected data, there are only 21 groups (accounting for 3.3%) with absolute error exceeding 10 seconds. The maximum absolute error (19.46 seconds) occurs when 17 passengers board, 19 passengers alight, and the carriage crowding level is 0.255.

In model II, 18.6% of absolute errors are less than 3 seconds. However, the absolute errors lager than 5 seconds account for 65.9% of total samples, with 31.9% and 6.7% exceeding 10 seconds and 20 seconds, respectively. The maximum absolute error in model II reaches up to 39 seconds, where 20 passengers board, 6 passengers alight, and the carriage crowding level is 0.956.

Further comparison reveals that the maximum absolute errors in both models occur when the numbers of boarding and alighting passengers are large. This is because among all the collected data, 70.9% and 85.4% refer to the case where boarding and alighting passengers are less than 10 people and 15 people, respectively. Meanwhile, these data play a vital role in the parameter regression of the dwell time estimation model. Consequently, both estimation models are less accurate when the numbers of boarding and alighting passengers are relatively large.
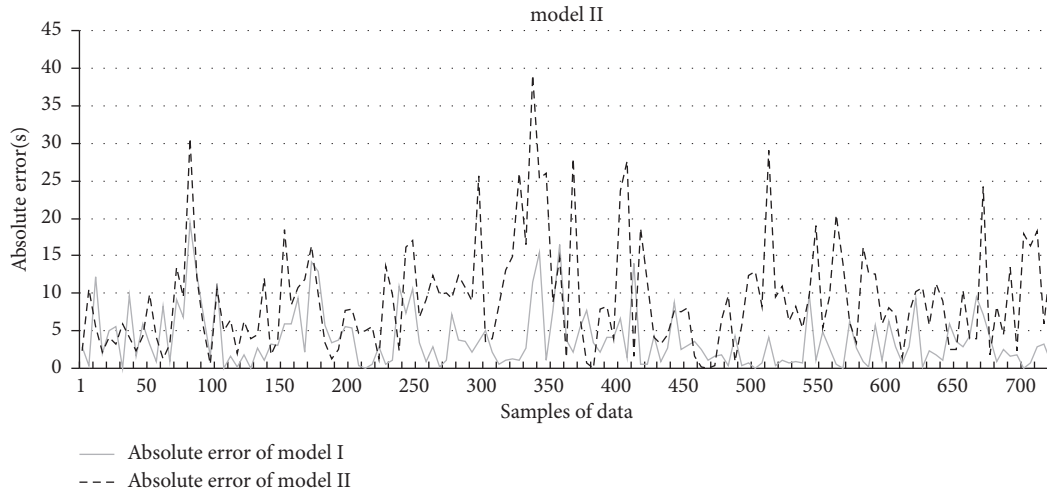
FIGURE 5: Comparison of the absolute errors between model I and model II.

TABLE 7: Comparison of MAE between model I and model II (unit: second).

Model	Boarding time	Alighting time	Dwell time
Model I	3.324	3.134	3.904
Model II	8.943	7.917	9.274

TABLE 8: Part of actual survey data of Route 8.

Data type	Values of the survey data					
Number of boarding passengers	2	2	5	5	8	8
Number of alighting passengers	6	7	8	7	0	0
Carriage crowing level	0.078	0.396	0.299	0.806	0.226	0.986
Actual dwell time (s)	9	13	14	21	21	34

4. Conclusions

The purpose of this research is to study the effect of bus crowding level on bus dwell time. The data is collected along two major bus routes in Harbin by manual data collection. Two statistical models are generated in this paper: one is developed based on the effect of both carriage crowding and passenger volume, while the second model only considers the effect of the number of boarding and alighting passengers on bus dwell time. By comparing these two models, we get some useful insights summarized as follows:

(1) The accuracy of model I is evidently higher than that of model II, with MAE reduced by 137.51%. It displays that the precision of bus dwell time model can be improved significantly by considering bus crowding level.

(2) Except for bus crowding level, bus dwell time is also influenced by the total amount of time used for opening and closing bus doors as well as passenger preparation. According to our analysis, this time is about 7 seconds.

(3) Our model also reveals that a 1% increase in the number of boarding (alighting) passengers will lead to 0.926% (0.848%) increase in boarding (alighting) time. Meanwhile, a 1% increase in crowding level can increase the boarding (alighting) time by 0.085% (0.092%).

In the future, the developed model would be tested on more bus routes. The impacts of other contributing factors can also be considered, such as the route overlap, the usage of mobile payment, and weather conditions. In addition, the developed model is also useful to improve the bus travel time prediction accuracy, which is important in dynamic bus scheduling.

Data Availability

The data supporting the model development of the study can be found in "Data Collection" section.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] X. Ma, Y.-J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders' travel patterns," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 1–12, 2013.

[2] Y. Bie, R. Tang, and L. Wang, "Bus scheduling of overlapping routes with multi-vehicle types based on passenger OD data," *IEEE Access*, vol. 8, pp. 1406–1415, 2019.

[3] X. Qu, M. Zhou, Y. Yu, C. T. Lin, and X. Wang, "Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: a reinforcement learning based approach," *Applied Energy*, vol. 257, Article ID 114030, 2020.

[4] M. Zhou, Y. Yu, and X. Qu, "Development of an efficient driving strategy for connected and automated vehicles at signalized intersections: a reinforcement learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 433–443, 2020.

[5] S. Liang, M. Ma, and S. He, "Multiobjective optimal formulations for bus fleet size of public transit under headway-based holding control," *Journal of Advanced Transportation*, vol. 2019, Article ID 2452348, 14 pages, 2019.

[6] D. Arbis, T. H. Rashidi, V. V. Dixit, and U. Vandebona, "Analysis and planning of bicycle parking for public transport stations," *International Journal of Sustainable Transportation*, vol. 10, no. 6, pp. 495–504, 2016.

[7] X. Ma, C. Liu, H. Wen, Y. Wang, and Y. J. Wu, "Understanding commuting patterns using transit smart card data," *Journal of Transport Geography*, vol. 58, pp. 135–145, 2017.

[8] Y. Bie, X. Xiong, Y. Yan, and X. Qu, "Dynamic headway control for high-frequency bus line based on speed guidance and intersection signal adjustment," *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 1, pp. 4–25, 2020.

[9] R. Rajbhandari, S. I. Chien, and J. R. Daniel, "Estimation of bus dwell times with automatic passenger counter information," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1841, no. 1, pp. 120–127, 2003.

[10] K. Kepaptsoglou and M. Karlaftis, "Transit route network design problem: review," *Journal of Transportation Engineering*, vol. 135, no. 8, pp. 491–505, 2009.

[11] T. K. Ojo, "Quality of public transport service: an integrative review and research agenda," *Transportation Letters*, vol. 11, no. 2, pp. 104–116, 2019.

[12] A. Ermagun, T. H. Rashidi, and A. Samimi, "A joint model for mode choice and escort decisions of school trips," *Transportmetrica A: Transport Science*, vol. 11, no. 3, pp. 270–289, 2015.

[13] M. Babaei, M. Rajabi-Bahaabadi, and A. Shariat-Mohaymany, "Estimation of travel time reliability in large-scale networks," *Transportation Letters*, vol. 8, no. 4, pp. 229–240, 2016.

[14] Y. Wang, Y. Bie, and Q. An, "Impacts of winter weather on bus travel time in cold regions: case study of Harbin, China," *Journal of Transportation Engineering, Part A: Systems*, vol. 144, no. 11, Article ID 05018001, 2018.

[15] Y. Bie, D. Wang, and H. Qi, "Prediction model of bus arrival time at signalized intersection using GPS data," *Journal of Transportation Engineering*, vol. 138, no. 1, pp. 12–20, 2012.

[16] Y. Bie, X. Gong, and Z. Liu, "Time of day intervals partition for bus schedule using GPS data," *Transportation Research Part C: Emerging Technologies*, vol. 60, pp. 443–456, 2015.

[17] Y. Lin, X. Yang, and N. Zou, "Passive transit signal priority for high transit demand: model formulation and strategy selection," *Transportation Letters*, vol. 11, no. 3, pp. 119–129, 2019.

[18] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.

[19] S. Wang and X. Qu, "Station choice for Australian commuter rail lines: equilibrium and optimal fare design," *European Journal of Operational Research*, vol. 258, no. 1, pp. 144–154, 2017.

[20] S. Wang, X. Qu, and Y. Yang, "Estimation of the perceived value of transit time for containerized cargoes," *Transportation Research Part A: Policy and Practice*, vol. 78, pp. 298–308, 2015.

[21] Z. Liu, Y. Yan, X. Qu, and Y. Zhang, "Bus stop-skipping scheme with random travel time," *Transportation Research Part C: Emerging Technologies*, vol. 35, pp. 46–56, 2013.

[22] M. T. Li, F. Zhao, L. F. Chow, H. Zhang, and S. C. Li, "Simulation model for estimating bus dwell time by simultaneously considering numbers of disembarking and boarding passengers," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1971, no. 1, pp. 59–65, 2006.

[23] F. Li, Z. Duan, and D. Yang, "Dwell time estimation models for bus rapid transit stations," *Journal of Modern Transportation*, vol. 20, no. 3, pp. 168–177, 2012.

[24] E. M. González, M. G. Romana, and O. M. Á. M., "Bus dwell-time model of main urban route stops," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2274, no. 1, pp. 126–134, 2012.

[25] S. Rashidi and P. Ranjitkar, "Estimation of bus dwell time using univariate time series models," *Journal of Advanced Transportation*, vol. 49, no. 1, pp. 139–152, 2015.

[26] S. Rashidi and P. Ranjitkar, "Bus dwell time modeling using gene expression programming," *Computer-Aided Civil and Infrastructure Engineering*, vol. 30, no. 6, pp. 478–489, 2015.

[27] A. Kathuria, M. Parida, C. Sekhar, and M. Pathak, "Examining bus lost time dynamics for a bus rapid transit station," *Journal of Public Transportation*, vol. 19, no. 2, pp. 168–182, 2016.

[28] X. Liu, Y. Yang, M. Meng, and A. Rau, "Impact of different bus stop designs on bus operating time components," *Journal of Public Transportation*, vol. 20, no. 1, pp. 104–118, 2017.

[29] R. P. Guenthner and K. Hamat, "Transit dwell time under complex fare structure," *Journal of Transportation Engineering*, vol. 114, no. 3, pp. 367–379, 1988.

[30] J. C. Levine and G. W. Torng, "Dwell-time effects of low-floor bus design," *Journal of Transportation Engineering*, vol. 120, no. 6, pp. 914–929, 1994.

[31] M. Milkovits, "Modeling the factors affecting bus stop dwell time: use of automatic passenger counting, automatic fare counting, and automatic vehicle location data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2072, no. 1, pp. 125–130, 2008.

[32] G. Currie, A. Delbosc, and J. Reynolds, "Modeling dwell time for streetcars in Melbourne, Australia, and Toronto, Canada," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2275, no. 1, pp. 22–29, 2016.

[33] S. Jaiswal, J. Bunker, and L. Ferreira, "Influence of platform walking on BRT station bus dwell time estimation: Australian analysis," *Journal of Transportation Engineering*, vol. 136, no. 12, pp. 1173–1179, 2010.

[34] R. Fernández, P. Zegers, G. Weber, and N. Tyler, "Influence of platform height, door width, and fare collection on bus dwell time," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2143, no. 1, pp. 59–66, 2010.

[35] A. Tirachini, "Bus dwell time: the effect of different fare collection systems, bus floor level and age of passengers," *Transportmetrica A: Transport Science*, vol. 9, no. 1, pp. 28–49, 2013.

[36] Q. Meng and X. Qu, "Bus dwell time estimation at bus bays: a probabilistic approach," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 61–71, 2013.

[37] D. Katz and L. A. Garrow, "The impact of bus door crowding on operations and safety," *Journal of Public Transportation*, vol. 15, no. 2, pp. 71–93, 2012.

[38] G. Fletcher and A. El-Geneidy, "Effects of fare payment types and crowding on dwell time," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2351, no. 1, pp. 124–132, 2013.