Received October 7, 2019, accepted October 29, 2019, date of publication November 5, 2019, date of current version November 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2951604

Spatial-Temporal Correlation Prediction Modeling of Origin-Destination Passenger Flow Under Urban Rail Transit Emergency Conditions

CHEN LI[1], JIANLING HUANG[1,2,3], BO WANG[2,3], YUYANG ZHOU[1], YUNYUN BAI[2,3], AND YANYAN CHEN[1]

[1]Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, China
[2]Beijing Transportation Information Center, Beijing 100161, China
[3]Beijing Key Laboratory for Comprehensive Traffic Operation Monitoring and Service, Beijing 100161, China

Corresponding author: Bo Wang (bowang@bjtu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1600900, in part by the Natural Science Foundation of Beijing under Grant 8171003, and in part by the Beijing Natural Science Foundation Program under Grant L181002.

ABSTRACT Establishing a passenger flow prediction mechanism is necessary for quickly evacuating many passengers in an emergency, which can improve the service quality of urban rail transit (URT). To effectively forecast origin-destination (OD) passenger flows in URT under emergency conditions, 35-day automatic fare collection (AFC) data are used for a statistical analysis of the time, location and passenger flow aspects. The influence range of the OD passenger flow during an emergency is determined by analyzing the degree of passenger flow fluctuation. Considering the time period of an emergency occurrence and its continuous influence, this paper also studies the influence of an emergency occurring at a station, a section between two stations or a section across several stations. A spatial-temporal correlation prediction model of OD passenger flow based on nonlinear regression is constructed by introducing the concept of passenger flow spatial-temporal influencing parameters. According to the characteristics of URT lines, a passenger flow prediction algorithm is proposed to predict the OD passenger flow for different line categories for an emergency. A real typical emergency involving the Beijing urban rail transit (BURT) system in 2017 is analyzed to verify the effectiveness of the proposed model. The results show that this model can effectively predict OD passenger flow in a URT system during an emergency, which provides basic support for the evacuation of passengers.

INDEX TERMS Urban rail transit (URT) emergency, origin-destination (OD) passenger flow prediction, spatial-temporal correlation, nonlinear regression, automatic fare collection (AFC) data.

I. INTRODUCTION

Urban rail transit (URT) has attracted more passengers as a main kind of transport mode because of its rapidity and punctuality. At present, URT is developing rapidly with the continuous expansion of the passenger flow. By the end of 2018, 493 cities in 72 countries and regions had opened a URT system, with a total operating mileage exceeding 26100 km [1]. As one of the countries with the fastest development of URT, the mileage of URT in China accounted for approximately 24% of the world's URT in 2018 according to data released by the China Urban Rail Transit Association. The operating mileage of Beijing urban rail transit (BURT) reached 637 kilometers. The daily passenger flow was more than 13 million in 2018, with the maximum section load rate already exceeding 140%, compared with 3.62 million in 2008. Because of the increasing number of passengers traveling by URT, many URT systems are overloaded, which has resulted in an increase in the frequency and types of emergencies. In total, 1500 emergencies that affected passenger travel occurred in the BURT system throughout 2018. One incident lasting for more than 10 minutes or causing a delay of at least 10 minutes occurred every 5 days in Singapore's mass rapid transit system from 2012 to 2016 [2]. Once an emergency occurs, it will affect the normal operation of all trains on the line. If not dealt with in time, it will result in

The associate editor coordinating the review of this manuscript and approving it for publication was Jesus Felez.

a decline in the transport capacity or even paralysis of the URT network. For example, an emergency caused by a signal failure occurred in August 2018 in the BURT system, which caused 40 trains to stop running, and the maximum delay for passengers was 49 minutes.

The causes and classifications of emergencies are numerous [3]-[5]. In general, URT mass passenger flows can be divided into predictable mass passenger flows and unpredictable emergency passenger flows. Passengers can be evacuated by planning ahead for the predictable situations such as holidays and large-scale activities. However, URT system failures, natural disasters, fires and other unpredictable events occur suddenly, and the associated responses are relatively difficult and cumbersome. The priority is to ensure passenger safety under natural disaster and fire conditions. In contrast, URT system equipment faults during daily operation are the most common events, and the proportions of signal failure and vehicle failure were 59.21% and 26.86%, respectively, for BURT emergencies in one year, which were the main types of equipment failures. In addition, the pressure to handle emergencies is relatively high when an emergency occurs during peak hours. Automatic fare collection (AFC) data provide accurate data support for URT passenger flow statistics and analysis. By using AFC data, passenger flow between any two stations in the URT network can be extracted, namely, the OD passenger flow. Therefore, the analysis subject in this paper is OD passenger flow under unpredictable emergencies, which results from signal and vehicle equipment failures during rush hour.

After an emergency, the original operation mechanism of the URT system appears to be temporary disorder. Passengers have unnecessary panic, and this anxiety spreads rapidly [6], [7], which will lead to changes in the original travel route. Passenger flow evacuation becomes extremely difficult in an emergency. Therefore, it is urgent to establish an early-warning system for mass passenger flow to eliminate or reduce the impact of an emergency on passengers and train operations.

The purpose of this study is to predict changes in OD passenger flow by exploring the spatial-temporal distribution rules of passenger flow after a URT emergency, which provides effective support for passenger flow evacuation. Based on 35-day AFC data in the BURT system, OD passenger flow prediction in cases of emergency is studied. The influence range of OD passenger flow in an emergency is proposed, and a spatial-temporal correlation prediction model of OD passenger flow is established in combination with the spatial-temporal influencing parameters related to passenger flow. Then, an OD passenger flow prediction algorithm is proposed to predict different line categories during an emergency. With the help of AFC data, this paper adopts a data-driven regression method, which considers the affected degree of passenger flow during emergencies spanning different time periods and in different locations. OD passenger flow prediction is the focus of this paper, and the research object is the fluctuation of OD passenger flow in an emergency, which can be obtained from AFC data. Therefore, the individual behavioral choices and travel routes in the URT system need no additional analysis, but passengers' choices can be reflected indirectly by changes in the amount of AFC data.

The rest of this paper is organized as follows. Section II introduces the related research works. In section III, the OD passenger flow prediction model is built. In section IV, the prediction algorithm for emergencies is proposed. Section V analyzes the case of the BURT system and evaluates the predicted results. Finally, the conclusion and discussion are given.

II. RELATED WORKS

In order to discuss OD passenger flow prediction under URT emergency conditions, in related works, we mainly concentrate on the studies regarding the influence range of an emergency, correlation analysis, and passenger flow prediction.

A. INFLUENCE RANGE OF AN EMERGENCY

After a breakdown in traffic, passenger flow evacuation becomes the primary concern [8], [9]. However, it is helpful to make passenger flow forecasts and evacuations more accurate to clarify the effect of an emergency. Tympakianaki *et al.* [10] considered the identification of impacted areas and the evaluation of impacts on network performance while exploring the multi-mode impact of traffic network interruption with heterogeneous data sources. Wen *et al.* [2] formulated a maximum survivability-minimum recovery time approach to determine the number of affected passengers in the rail network during a disruption, the number of passengers who need to be transferred to alternate transport modes, and the recovery duration. Sun *et al.* [11] proposed an efficient Bayesian method to identify the affected stations during a disruption in a URT system. These studies investigate the influence range of emergencies from one or two perspectives of passenger flow, space or time. Zhang *et al*. [12] considered the length of evacuation route, the time of evacuation process and the density of pedestrian flow, but they mainly analyzes the factors affecting the evacuation efficiency of passenger flow after an emergency, without considering the occurrence period of the emergency.

After an emergency, it is important to study the influence range of passenger flow, but the degree of temporal and spatial influences of an emergency during different periods is also closely related to passenger flow evacuation, which is comprehensively studied in this paper.

B. CORRELATION ANALYSIS

The factor correlation analysis method has been widely used in road traffic passenger flow analysis. Xu *et al.* [13] adopted the association rule mining technique to identify sets of crash contributory factors in serious casualty crashes. Chen and Deng [14] presents a hierarchical clustering analysis to explore the relationship between information use, social networks, and cooperation consciousness of individual travelers. Das *et al.* [15] used Multiple Correspondence Analysis to analyze Run-Off-Road crashes and measures important contributing factors and their degree of association. However, spatial-temporal correlation factors have a significant impact on traffic accidents and residents' travel. Pan *et al.* [16] forecasted short-term traffic states based on the spatial-temporal correlation. Wang *et al.* [17] designed a time-series constraint and an adaptive Laplacian regularization spatial constraint to explore the temporal evolvement characteristics and the spatial similarity of road links.

URT is different from road traffic, in which how passengers move depends on the URT train. Once a train is interrupted, passengers on the train will be affected. After a URT emergency ends, passenger flow may increase in some stations or lines. When conducting a correlation analysis of the main factors affecting OD passenger flow in a URT emergency, it is necessary to consider the characteristics of the URT system. Wan *et al.* [18] used logistic regression to analyze the factors affecting the occurrence of URT accidents based on questionnaire data, which data might be affected by the sample and human factors. It is a good choice to use AFC data to analyze the correlation of influencing factors when an emergency occurs in the URT system.

C. PASSENGER FLOW PREDICTION

Existing literature on URT passenger flow prediction estimates future traffic demands through four-step transportation planning model or regression techniques [19], neural network [20], deep learning [21] and so on. Most of these methods are applicable to the regular passenger flow prediction with small fluctuations, while the distribution curve of a sudden large amount of passenger flow is an obvious abrupt change compared with normal passenger flow [12]. Moreover, considering the contingency and randomness of an emergency occurrence, the probability of an emergency of the same type at the same time and location is infinitesimal, which results in insufficient prior information for neural network and support vector machine. Therefore, how to forecast passenger flow in emergency remains to be further explored.

Many experts and scholars have explored the passenger flow prediction of URT under abnormal conditions. Li *et al.* [22] proposed a novel multiscale radial basis function network to forecast irregular fluctuations in URT passenger flow. Ni *et al.* [23] proposed a model based on the fusion of support vector regression and a long short-term memory neural network to predict short-term abnormal passenger flow in the URT system. Guo *et al.* [24] forecasted URT passenger flow under event occurrences with social media. An *et al.* [25] proposed a novel fuzzy-based convolutional neural network to predict traffic flow with uncertain traffic accident information. These studies are mainly aimed at the prediction of inbound and outbound passenger flow under abnormal conditions, and there are relatively few studies on OD passenger flow prediction in emergencies. OD passenger flow reflects the degree of mutual attraction between stations. The inbound volume, outbound volume, and transfer volume can be obtained by passenger distribution in the case that OD passenger flow is determined. Therefore, OD passenger flow prediction is necessary and very important. Liu *et al.* [26] predicted the OD passenger flow of URT in an emergency based on the results of passenger choice behavior modeling. However, the data for this model mainly relied on a questionnaire survey, which might be greatly affected by the sample and human factors.

In the case of relatively abundant historical data, this paper adopts a regression analysis to establish the relationship between passenger flow and influencing factors, which is more suitable for OD passenger flow prediction related to an emergency.

III. MODEL DEVELOPMENT

A. ASSUMPTIONS

Considering the influence of passenger flow characteristics and emergency measures under emergencies comprehensively, OD trips after an emergency can be divided into four categories: OD trips for which the origin station or destination station are on the emergency line, OD trips with a travel route going through the emergency line, OD trips with a travel route going through the line that can transfer to the emergency line, and other OD trips. Moreover, OD trips with a travel route can transfer to the emergency line should also be affected by an emergency. However, for the URT network, these OD trips can be divided into two cases. One is that these passengers transfer to the emergency line and get out of the station on the emergency line, which corresponds to the first category. The other is the OD trips in which the travel route passes through the emergency line, which corresponds to the second category. The proportions of OD trips are calculated as follows.

$$f(P_i)=\begin{cases}\dfrac{\mathrm{card}(P_i)}{\mathrm{card}(Q)}, & i=1,2,3\\ 1-\dfrac{\mathrm{card}(P_1)+\mathrm{card}(P_2)+\mathrm{card}(P_3)}{\mathrm{card}(Q)}, & i=4\end{cases} \tag{1}$$

where Q is the set of OD trips within the influence range of OD passenger flow in an emergency, P_i is the set of OD trips of different categories, card (Q) and card (P_i) are the number of elements in sets Q and P_i, respectively. When $i = 1$, the travel route for which the origin or destination station is on the emergency line. When $i = 2$, the travel route is passing through the emergency line. When $i = 3$, the travel route is passing through the line that can transfer to the emergency line.

Effective data screening and statistics are carried out for two BURT emergencies, as shown in Table 1. Event 1 occurred on Line 10 in November 2017, which caused 33 trains to stop running. In order to reduce the impact of passenger flow, the operating enterprise took measures to skip the transfer station that could transfer to the emergency line. For event 2, it happened on Line 1 in August 2017, resulting in 20 trains to stop running. However, considering the overall operating efficiency, the operating enterprises did not take emergency measures.

TABLE 1. Ratios of OD pairs for different categories for two emergencies in Beijing urban rail transit system.

$f(i)$	Event 1	Event 2
$f(P_1)$	60.49%	65.55%
$f(P_2)$	0.00%	0.00%
$f(P_3)$	16.89%	0.00%
$f(P_4)$	22.62%	34.45%

For the two emergencies, the proportion for set P_1 is the largest, which is the OD pair set with a larger influence on the emergency. By contrast, the proportion for set P_2 is 0, which can be ignored. The proportion of OD trips in set P_3 is related to emergency measures taken by the operating enterprises. For event 1, the operating enterprise took measures to let the train on the adjacent lines to the emergency line skip a transfer station, so passengers could not transfer between the emergency line and the adjacent line. If the operation enterprise takes emergency measures that affect the passenger flow of other lines, it can be considered another mass passenger flow event, which is event superposition. The OD trips in set P_4 also have a certain proportion, but these trips are usually caused by other emergencies in the same period or fluctuations in passenger flow caused by transportation hubs around URT stations, and OD passenger flow changes have nothing to do with this emergency, so it needs to be excluded.

To forecast OD passenger flow affected by an emergency more accurately, we first make the following assumptions throughout this paper.

Assumption 1: The operating enterprise does not take emergency measures in the URT network after an emergency occurs.

Assumption 2: OD trips of the origin station or destination station on the emergency line would be affected by the emergency, while other OD trips are not affected.

Assumption 3: An emergency has only a certain impact on the areas in which passenger flow fluctuates greatly. For unaffected passenger flows, the trend in OD passenger flow conforms to the distribution of passenger flow in the same historical period.

B. INFLUENCE RANGE OF AN EMERGENCY

To study the influence range of an emergency, we elaborate on and analyze three aspects: time, space, and passenger flow.

1) TIME INFLUENCE OF AN EMERGENCY

Once an emergency occurs, it needs to be dealt with promptly. Generally, the time granularity of passenger flow research is 5 minutes, 15 minutes, 30 minutes or 1 hour [27]. Under normal circumstances, OD passenger flow with a granularity of 5 minutes is easily affected by unstable travel times, route changes of passengers and other factors. For example, if one passenger leaves home 10 minutes earlier, he or she will arrive at the destination station approximately 10 minutes earlier. At the same time, an OD passenger flow of 5 minutes is relatively small, and small data fluctuations will cause errors in the data analysis. When an emergency occurs, it is difficult to determine whether fluctuations in OD passenger flow under a granularity of 5 minutes is caused by the change in passenger travel behaviors or the emergency. Therefore, to improve the accuracy of OD passenger flow analysis and prediction for an emergency, this paper chooses 15-minute granularity for the analysis.

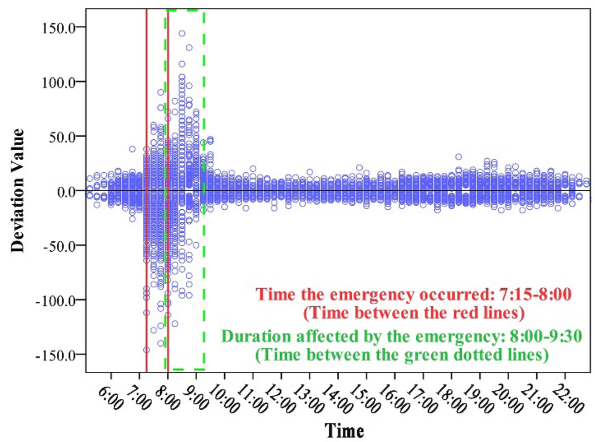


FIGURE 1. OD passenger flow deviation value distribution in an emergency.

Fig. 1 shows the distribution of OD passenger flow over time after an emergency on Line 10 of the BURT system. In each period, OD passenger flow has slight fluctuations under normal circumstances, so data cleaning and screening should be carried out in the passenger flow analysis. OD passenger flow fluctuations can be described by OD passenger flow deviation value, which is the difference between the OD passenger flow volume after an emergency occurs and the normal passenger flow.

This emergency occurred from 7:15 to 8:00, which corresponded to the time period between the red lines in Fig. 1. OD passenger flow was still affected after emergency, as shown by the area between the green dotted lines. Thus, the occurring period and the duration of influence should be considered when forecasting OD passenger flow. When the deviation value of OD passenger flow is basically 0, it can be considered that the continuing influence of OD passenger flow ends. The corresponding moment is the ending time of the continuous influence of the emergency.

The occurrence time of an emergency is the starting and ending time published by the operating enterprise. The continuous influence time of an emergency is related to the type of emergency, emergency-occurring period, and the URT network. As shown in Fig. 2, T_1 and T_e are the starting and ending times of the emergency, T_c is the ending time of when the emergency affects the URT network, t_e is the time interval of the emergency, t_c is the time interval during which the emergency continues to affect the URT network. The units of t_e and t_c are minute.

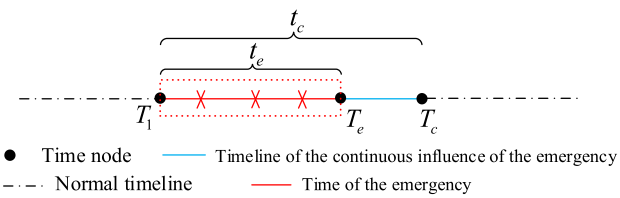


FIGURE 2. Diagram of the influence time for passenger flow in an emergency.

To describe the relationship between the OD passenger flow deviation degree and time, the concept of a passenger flow time influence parameter is proposed. Define the passenger flow time influence parameter φ to represent each 15-minute stage within the time period affected by emergency. The first 15 minutes of an emergency can be denoted by 1, the second 15 minutes can be denoted by 2, and so on. The last 15 minutes before the end of the emergency can be denoted by $\left\lceil \frac{t_e}{15} \right\rceil$.

After the end of an emergency, the continuous influence of passenger flow should be analyzed according to the change in passenger flow, which can be determined by the function $J(t)$.

Therefore, within the time influence range of an emergency, the passenger flow time influence parameter can be expressed as $\varphi = 1, 2, \ldots, \left\lceil \frac{t_e}{15} \right\rceil, \left\lceil \frac{t_e}{15} \right\rceil + 1, \ldots, J(t)$. The value of $J(t)$ can be calculated by formula (2).

$$J(t) = \begin{cases} \left\lceil \dfrac{t_c}{15} \right\rceil, & N_{t,o\times d}^{emergency} \neq \phi \\ & \text{and } N_{t+1,o\times d}^{emergency} = \phi \\ \left\lceil \dfrac{t_c}{15} \right\rceil + 1, & \text{others} \end{cases} \tag{2}$$

where ϕ is an empty set; $N_{t,o\times d}^{emergency}$ is the label matrix of the affected OD pair during time period t, in which the number of rows and columns of the label matrix is the number of stations in a URT network, namely the matrix $o \times d$. If there is deviation for any OD pairs, the corresponding element in the set $N_{t,o\times d}^{emergency}$ is 1, otherwise it is 0. When there is one or more elements are 1 in the set $N_{t,o\times d}^{emergency}$, and all the elements are 0 in set $N_{t+1,o\times d}^{emergency}$, it is $J(t) = \left\lceil \frac{t_c}{15} \right\rceil$. Otherwise, it is $J(t) = \left\lceil \frac{t_c}{15} \right\rceil + 1$.

2) SPATIAL INFLUENCE OF AN EMERGENCY

After an emergency occurs, it is unevenly distributed for OD passenger flow for which the origin station or destination station was on the emergency line. Fig. 3 shows the deviation rate distribution of OD passenger flow for which the origin stations were on Line 10 when an emergency occurs on Line 10 in the BURT system.

The sections between the red straight lines were the locations where this emergency occurred, and the OD passenger flow, which affected stations within the green lines as the origin station, was affected by the emergency. Therefore, the influence of spatial location should be taken into account in the prediction of OD passenger flow.

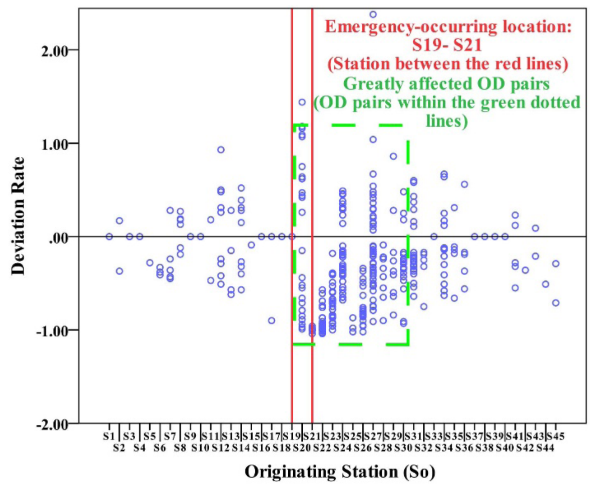


FIGURE 3. Deviation rate distribution of OD passenger flow for the original station on Line 10 after an emergency.

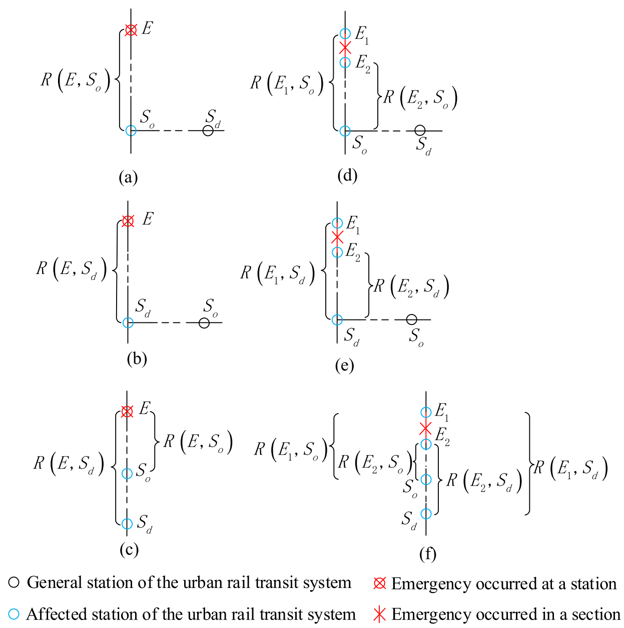


FIGURE 4. Distance relationship between emergency-occurring locations and OD travel.

The spatial influence of the emergency can be reflected by the distance between the emergency-occurring location and the origin or destination station of one trip. The location of an emergency can be a station, a section between two stations, or a section across several stations.

The distance between an emergency-occurring location and OD travel is shown in Fig. 4. S_o and S_d are the origin station and destination station, respectively.

When an emergency occurs at a station, as shown in Fig. 4 (a), (b), and (c), then

$$R = \begin{cases} R(E, S_o), & S_o \in \mathrm{K}, S_d \notin \mathrm{K} \\ R(E, S_d), & S_d \in \mathrm{K}, S_o \notin \mathrm{K} \\ \min\{R(E, S_o), R(E, S_d)\}, S_o \in \mathrm{K}, & S_d \in \mathrm{K} \end{cases} \tag{3}$$

When an emergency occurs in a section between two stations, as shown in Fig. 4 (d), (e), and (f), then

$$R=\begin{cases}\min\{R(E_1,S_o),R(E_2,S_o)\}, & S_o\in K, S_d\notin K\\ \min\{R(E_1,S_d),R(E_2,S_d)\}, & S_d\in K, S_o\notin K\\ \min\{R(E_1,S_o),R(E_1,S_d), & \\ R(E_2,S_o),R(E_2,S_d)\}, S_o\in K, & S_d\in K\end{cases}\tag{4}$$

When an emergency occurs in a section across several stations, the calculation of the distance is similar to that for an emergency occurring in a section between two stations. where K is the set of stations on the emergency line; R is the shortest distance between the origin or destination station to an emergency-occurring location; and $R(a,b)$ is the shortest distance between station a and b, where $a = E, E_1, E_2$, $b = S_o, S_d$.

3) INFLUENCE RANGE OF OD PASSENGER FLOW

After the occurrence of an emergency, passenger flow will have different deviations from the historical value in the same period. The OD passenger flow deviation rate is proposed, which is the ratio of the OD passenger flow deviation value after an emergency to the normal OD passenger flow. The relationships among the average value of the OD passenger flow, the OD passenger flow deviation value and the OD passenger flow deviation rate during the historical contemporaneous passenger flow are analyzed, as shown in Fig. 5.

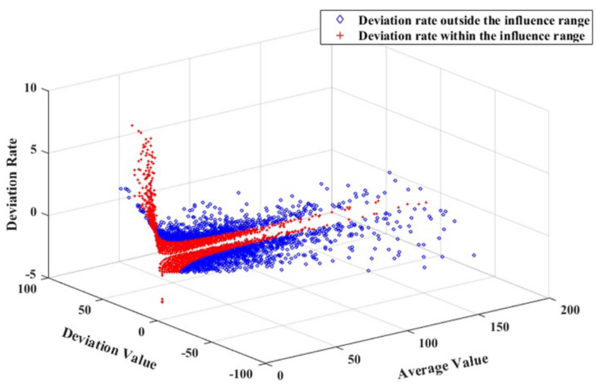


FIGURE 5. Distribution of the OD passenger flow deviation rate, deviation value, and average volume.

The OD passenger flow marked in red in Fig. 5 is within the influence range of the OD passenger flow in an emergency. This paper presents a method to identify the influence range of OD passenger flow during an emergency. In combination with the degree of OD passenger flow fluctuation in an emergency, the OD deviation value and deviation rate are used to define the influence range of passenger flow.

$$x_{t,od}^{normal} = \frac{1}{g}\times\sum_{j=1}^{g} x_{j,t,od}^{normal}, \quad j=1,2\cdots,g \tag{5}$$

$$x_{t,od}^{deviation} = x_{t,od}^{emergency} - x_{t,od}^{normal} \tag{6}$$

$$y_{t,od}^{deviation} = \frac{x_{t,od}^{emergency}}{x_{t,od}^{normal}} - 1 \tag{7}$$

where $x_{j,t,od}^{normal}$ is the OD passenger flow in period t of the jth normal day; $x_{t,od}^{normal}$ is the average OD passenger flow in period t; $x_{t,od}^{emergency}$ is the OD passenger flow of an emergency day during period t; $x_{t,od}^{deviation}$ is the OD passenger flow deviation value during period t; $y_{t,od}^{deviation}$ is the OD passenger flow deviation rate during period t; and g is number of normal days during the same historical period. When $-1 \le y_{t,od}^{deviation} < 0$, some passengers leave the URT system and choose other urban traffic modes to travel or abort this travel. When $y_{t,od}^{deviation} > 0$, part of the passenger flow transfers to the OD pair.

The range of OD passenger flow affected by an emergency during period t, which is set to Γ_t, is jointly constrained by $x_{t,od}^{normal}$, $x_{t,od}^{deviation}$, and $y_{t,od}^{deviation}$. For trips with a small OD passenger flow, slight fluctuations in the passenger flow will show a larger degree of change, so trips for which the historical OD passenger flow is smaller should be removed. Similarly, trips with smaller deviation values and deviation rates should also be removed.

$$\begin{cases} x_{t,od}^{normal} \ge \theta \\ abs(x_{t,od}^{deviation}) \ge abs\left(x_{\min}^{deviation}\right) \\ abs(y_{t,od}^{deviation}) \ge abs\left(y_{\min}^{deviation}\right)\end{cases}\tag{8}$$

where θ is the OD passenger flow lower limit parameter, which depends on the historical OD passenger flow; $x_{\min}^{deviation}$ is the minimum deviation value of the OD passenger flow within the influence range; and $y_{\min}^{deviation}$ is the minimum deviation rate of the OD passenger flow within the influence range. The values of $x_{\min}^{deviation}$ and $y_{\min}^{deviation}$ depend on the historical normal OD passenger flow and the OD passenger flow in an emergency.

For OD pairs outside the influence range, the OD passenger flow conforms to the historical normal daily trend of the same period; that is, the deviation rate is 0. The label for the OD pairs within the influence range is 1, while the label for the OD pairs outside the influence range is 0.

$$n_{t,o\times d}^{emergency} = \begin{cases} 1, & \left\{x_{t,od}^{normal}, x_{t,od}^{\mathrm{deviation}}, y_{t,od}^{\mathrm{deviation}}\right\} \in \Gamma_t \\ 0, & others\end{cases}\tag{9}$$

where $n_{t,o\times d}^{emergency}$ is the label of the OD pairs, which is the element of $N_{t,o\times d}^{emergency}$.

C. PASSENGER FLOW PREDICTION MODEL

Factors influencing passenger travel behavior in an emergency include passenger attributes, event time, travel time, travel cost, and transfer convenience [28]-[30]. Based on the characteristics and rules of OD passenger flow in an emergency, the factors affecting OD passenger flow are summarized, which mainly include the influence time, the emergency-occurring location, the inbound and outbound volume, and the travel distance.

The influencing time of an emergency includes the occurrence period and the continuous influence period, which can be expressed by the passenger flow time influence parameter. Emergency-occurring locations mainly include three types of stations, a section, and sections. Passenger flow, which is into the origin stations and out of the destination stations, has a great impact on OD travel. When an emergency occurs, the stations with a large amount of inbound and outbound volume will bear great pressure in passenger flow evacuation. At the same time, the travel distance directly affects the travel choice of passengers. The shortest OD travel distance is used as the travel distance to study the correlation between passenger travel distance and OD passenger flow deviation rate.

Considering that the inbound volume of the origin station and the outbound volume of the destination station correspond to OD passenger flow, the OD passenger flow deviation rate in an emergency can be discussed. OD passenger flow in an emergency can be predicted by combining the normal daily OD passenger flow and deviation rate.

A correlation analysis between the OD passenger flow deviation rate and the time influence parameter (φ), the distance from the origin or destination station to the emergency-occurring location (R), the distance of the shortest route between the origin and destination station (R_{od}) is used in this study. Data screening and cleaning are conducted for an emergency in the BURT network in 2017. Take a BURT emergency that occurred in August 2017 as an example, the effective statistical sample of OD pairs is 3486. The correlation analysis results are shown in Table 2.

TABLE 2. Results of the correlation analysis.

Influencing Factor	Correlation Parameter	Significance Value (2-tailed)
φ	0.596	0.000
R	0.113	0.000
R_{od}	-0.073	0.000

From the results of the correlation analysis, we find that all the significance values are less than 0.05, which suggests that these factors are correlated. However, the correlation coefficients of R with $x_{t,od}^{deviation}$, and R_{od} with $x_{t,od}^{deviation}$ are small, which means that their correlations are less significant.

To more accurately describe the relationship between spatial factors and the OD passenger flow deviation, the concept of a passenger flow spatial influence parameter is proposed to describe the degree to which the OD passenger flow affected by the spatial distance in an emergency.

$$\delta_t^{od} = \frac{R}{R_{od}} \tag{10}$$

where δ_t^{od} is the passenger flow spatial influence parameter.

The correlation analysis between the OD passenger flow deviation rate and the OD passenger flow spatial influence parameter from three perspectives (i.e., the whole time period, the occurrence time period and the continuous influence period of the emergency) are shown in Table 3.

TABLE 3. Correlation analysis of the passenger flow spatial influence parameter and the OD passenger flow deviation rate.

Influencing Factor	Whole Time Period	Occurrence Time	Duration
Correlation Coefficient	-0.362	0.618	-0.493
Significance Value (2-tailed)	0.000	0.000	0.000

The correlation significance values are less than 0.05. The correlation coefficients are -0.362, 0.618, and -0.493, respectively. The results show that the correlation coefficients are significantly improved after the time period is divided into an occurrence time period and a continuous influence period.

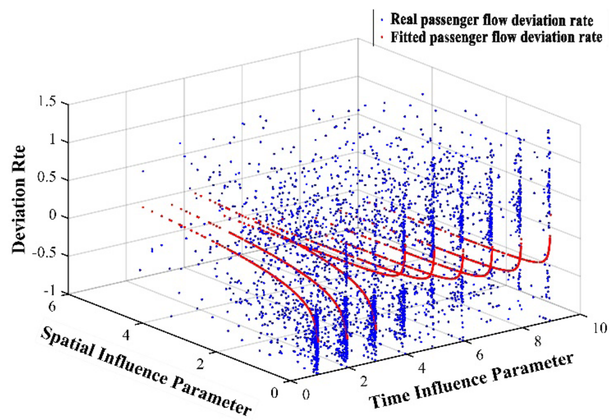


FIGURE 6. Distribution of the OD passenger flow deviation rate with the passenger flow spatial-temporal influence parameters.

The deviation rate distribution with the passenger flow spatial-temporal influence parameter is shown in Fig. 6. It is not easy to observe the variation trend in OD passenger flow in three-dimensional space.

The OD passenger flow deviation rate affected by different regions in the same period is averaged to obtain the average OD passenger flow deviation rate, which shows an obvious and stable trend over time, as shown in Fig. 7.

Without considering the influence of the spatial region, the impact form of an emergency conforms to the form of a sine function, which can be expressed by the following formula.

$$f(\varphi) = A \times \sin(B \times \varphi + C) + D \tag{11}$$

where A, B, C, and D are parameters.

The propagation of an emergency is not uniformly outward. The OD passenger flow deviation rate varies with different influence distances at the same time. The deviation rate distribution with the spatial passenger flow influence parameter is shown in Fig. 8.

From Fig. 8, we can see that for the emergency-occurring period, the OD passenger flow deviation rate follows a logarithmic distribution with the change in passenger flow spatial

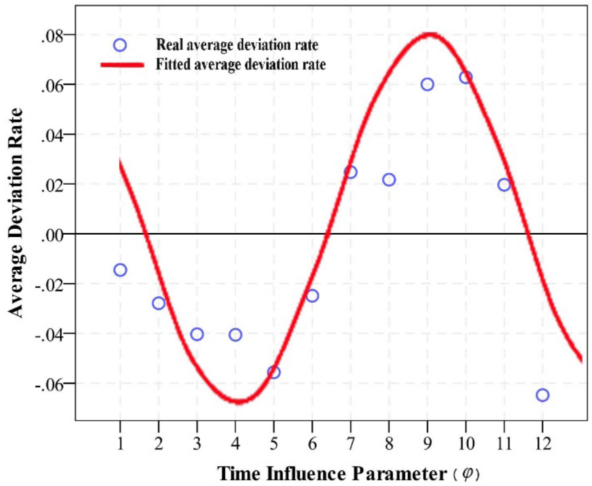


FIGURE 7. Time distribution of the OD passenger flow average deviation rate.

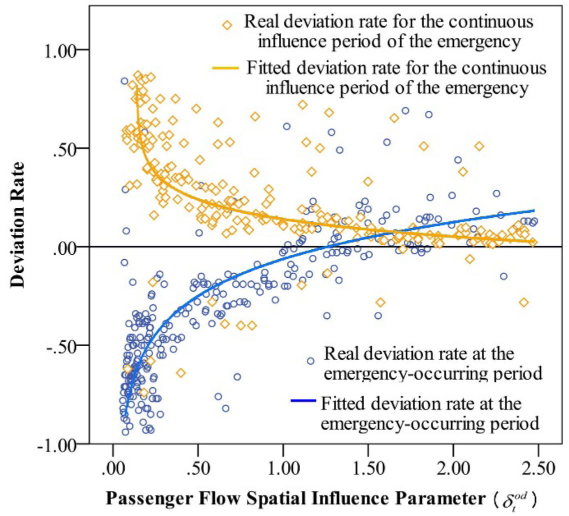


FIGURE 8. Spatial distribution of the OD passenger flow deviation rate.

influence parameters, while the change follows a power function distribution in the time period of the continuous emergency influence period. We can express these distributions with the following formula.

$$f\left(\delta_t^{od}\right) = \begin{cases} E \times \ln\left(\delta_t^{od}\right) + F, & t \in (T_1, T_e) \\ G \times \left(\delta_t^{od}\right)^H + I, & t \in (T_e, T_c) \end{cases} \tag{12}$$

where E, F, G, H, and I is are the parameters.

Considering the impact of passenger flow fluctuations such as transportation hubs and tourist areas in an emergency, there will be a small amount of abnormal data on the impact of the spatial passenger flow. In the regression analysis, this part of the data must be cleaned to improve the fitting accuracy.

From the perspective of the temporal and spatial correlations, the OD passenger flow prediction model in an emergency is constructed with the method of weighted combination, as shown in formulas (13) and (14).

$$y_{t,od}\left(\varphi, \delta_t^{od}\right) = \begin{cases} \lambda_1 \times \sin(\lambda_2 \times \varphi + \lambda_3) + \lambda_4 \times \ln\left(\delta_t^{od}\right) + \lambda_5, \\ \quad \varphi \in \left(1, \left\lceil \frac{t_e}{15} \right\rceil\right) \\ \mu_1 \times \sin(\mu_2 \times \varphi + \mu_3) + \mu_4 \times \left(\delta_t^{od}\right)^{\mu_5} + \mu_6, \\ \quad \varphi \in \left(\left\lceil \frac{t_e}{15} \right\rceil + 1, J(t)\right) \end{cases} \tag{13}$$

$$x_{t,od} = \begin{cases} \left(1 + y_{t,od}\left(\varphi, \delta_t^{od}\right)\right) \times x_{t,od}^{normal}, \\ \quad \left\{x_{t,od}^{normal}, x_{t,od}^{deviation}, y_{t,od}^{deviation}\right\} \in \Gamma \\ x_{t,od}^{normal}, \quad others \end{cases} \tag{14}$$

where $y_{t,od}\left(\varphi, \delta_t^{od}\right)$ is predicted value of the OD passenger flow deviation rate affected by an emergency during period t; $x_{t,od}$ is predicted value of the OD passenger flow during period t; and λ_m, μ_n are model parameters, where $m = 1, \cdots, 5$ and $n = 1, \cdots, 6$.

D. PASSENGER FLOW PREDICTION FOR DIFFERENT LINE TYPES

Due to land use properties, line attributes and other reasons, the passenger flow change trends for different line types in emergency are different, leading to differences in the parameters of the prediction model.

Let the set of all lines in the URT network be $\Omega = \{L_\eta\}$, where $\eta = 1, 2, 3, 4, 5$. The set of stations in urban areas is Z_1, while the set of stations in suburban areas is Z_2. The transfer station on lines m and n is H^{mn}. A line traveling through the urban area and suburban area of city is l_1, and the set of these stations is L_1. The stations on line l_1 are $S_{1,j}$, where $S_{1,j} \in l_1$. The line that starts in the suburbs and ends in the center of the city is l_2, and the set of these stations is L_2. The stations on line l_2 are $S_{2,j}$, where $S_{2,j} \in l_2$. The line that is in the suburbs is l_3, and the set of these stations is L_3. The stations on line l_3 are $S_{3,j}$, where $S_{3,j} \in l_3$. Loop lines are l_4, and the set of these lines is L_4. The stations on line l_4 are $S_{4,j}$, where $S_{4,j} \in l_4$. The set of the other remaining lines is L_5. Among them, $j = 1, \cdots, \rho, \cdots, \tau$, where $\tau \geq 2$ and $\tau \geq \rho$. Different types of URT lines are shown in Fig. 9.

Combined with Fig. 9, we give the conditions that each type of line needs to meet. The line throughout the urban area and suburban area of city should meet the following requirement:

$$L_1 = \left\{l_1 \left| \exists S_{1,j} \in Z_1, S_{1,j} \in Z_2, S_{1,1} \in Z_2, S_{1,\tau} \in Z_2, S_{1,j} \in l_1\right.\right\} \tag{15}$$

The line with one end in the suburbs and the other end in the center of the city needs to meet the following requirement:

$$L_2 = \left\{l_2 \left| \exists S_{2,j} \in Z_1, S_{2,j} \in Z_2, S_{2,1} \in Z_2, S_{2,\tau} \in Z_1, S_{2,j} \in l_2\right.\right\} \tag{16}$$

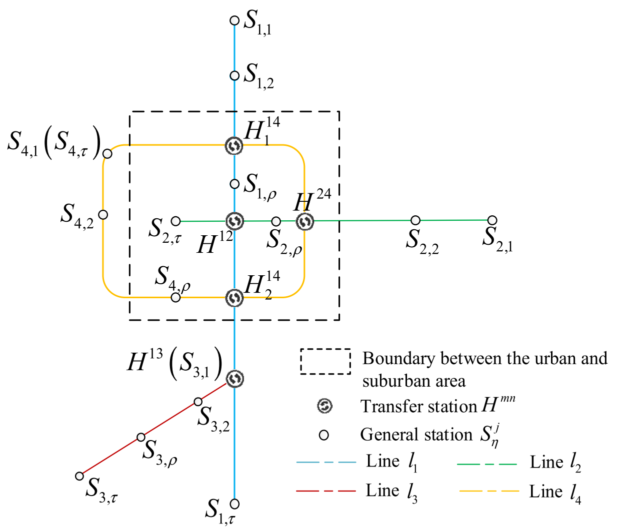


FIGURE 9. Diagram of the different types of lines.

The line that is in the suburbs needs to meet the following requirement:

$$L_3 = \left\{ l_3 \left| \forall S_{3,j} \in Z_2, S_{3,1} = T^{13} \right., S_{3,j} \in l_3 \right\} \tag{17}$$

Loop lines should meet the following requirement:

$$L_4 = \left\{ l_4 \left| \exists S_{4,1} = S_{4,\tau} \right., S_{4,j} \in l_4 \right\} \tag{18}$$

The remaining line set in the URT network is:

$$L_5 = \{ l_5 \, | l_5 \notin (L_1 \cup L_2 \cup L_3 \cup L_4) \} \tag{19}$$

According to the characteristics of the emergency line, the parameter set for the five types of URT lines is established. An OD passenger flow prediction model based on the emergency can be built. According to the characteristics of the emergency line, the appropriate model is retrieved from the set of prediction models, and the corresponding passenger flow is effectively predicted. Moreover, if the emergency-occurring line cannot be retrieved from the line type set Ω, the new model parameters can be quickly calculated and stored in line type set Ω.

IV. PASSENGER FLOW PREDICTION ALGORITHM

A prediction algorithm is proposed to forecast OD passenger flow in this paper. This algorithm also considers the characteristics of different lines, which can realize the prediction of OD passenger flow for multitype URT lines. URT line categories and corresponding passenger flow prediction models can be constantly enriched through feedback. The OD passenger flow prediction algorithm is as follows:

Step 1: Input the emergency information. The influence range of the OD passenger flow is identified according to formulas (5) - (8).

Step 2: Determine the emergency-occurring period and the continuous influence time period. Calculate the OD passenger flow time influence parameter combined with formula (2).

Step 3: Calculate the OD passenger flow spatial influence parameter according to formulas (3), (4), and (10). If an emergency-occurring line l_η is a member of set Ω, go to step 6; if not, go to step 4.

Step 4: Build the regression model and analyze the relationship between $y_{t,od}^{deviation}$, φ, and δ_t^{od}. Obtain the relationship $y_{t,od}^{deviation} \sim f\left(\varphi, \delta_t^{od}\right)$. Then, generate the parameter p, which includes λ_m and μ_n, from the OD passenger flow deviation rate prediction model.

Step 5: Store the line in set Ω, and store the parameter p in the parameter library P of the OD passenger flow deviation rate prediction model.

Step 6: According to the emergency-occurring line, retrieve the corresponding type of line in the set $\Omega = \left\{ L_\eta \right\}$. Calculate the predicted value of the OD passenger flow according to formulas (13) - (14).

V. CASE ANALYSIS

A. EMERGENCY INFORMATION

A real emergency of the BURT network in 2017 is selected for analysis. This emergency occurred between Tiantongyuan South Station and Lishuiqiao Station of Line 5 of the BURT system from 7:35 to 8:11, which was during the morning peak hour. The emergency resulted from a signal failure, causing 25 trains to stop running, 43 trains to be late, and 4 trains to make passengers exit the trains.

Line 5 is the typical line throughout the urban and suburban area of Beijing and services high-density residential and commercial areas. Once an emergency occurs, passengers will be greatly affected, especially during rush hour.

B. DATA PREPROCESSING

Let $g = 6$; that is, select the OD passenger flow data of 3 normal days in the same period before and after the day of the emergency. It should be noted that if the selected date is at a special time, such as a national holiday or a large activity day, or a fault occurs in the same period of the emergency, the OD passenger flow data of other similar days in the week will be selected.

To ensure the accuracy of the emergency identification, abnormal data with large deviations in the normal historical data should be cleaned. Assuming that the passenger flow during the same historical period follows a normal distribution, the single-sample K-S test and Q-Q figure are adopted for verification. The OD passenger flow during the morning and evening peak period from Shilihe Station to Guomao Station of the BURT system are selected for the normal distribution test, as shown in Table 4 and Fig. 10.

The asymptotic significance values of the single-sample K-S test in the morning and evening peak hour from Shilihe Station to Guomao Station are 0.605 and 0.755, respectively, which are much higher than the significance level of 0.05. Points on the Q-Q diagram are approximately distributed near the straight line, so the assumption that OD volume obeys a normal distribution in the same period is confirmed.

TABLE 4. Test results of a single-sample Kolmogorov-Smirnov test.

		Morning Peak Hour	Evening Peak Hour
N		32	32
Normal Parameters[a,b]	Mean	403.188	68.25
	Standard Deviation	84.2816	20.061
Most Extreme Differences	Absolute Value	0.135	0.119
	Positive Number	0.135	0.100
	Negative Number	-0.086	0.119
Kolmogorov-Smirnov Z		0.763	0.673
Asymptotic Significance Value(2-tailed)		0.605	0.755

a. The test distribution is normally distributed.

b. Calculated from data.

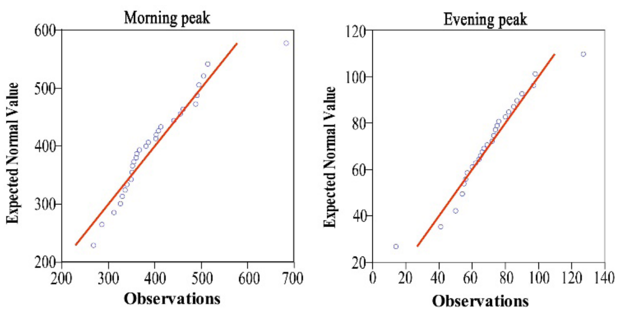


FIGURE 10. Normal Q-Q distribution of the OD volume in the morning peak and evening peak hour.

TABLE 5. Data processing parameters.

Parameter	Value
$x_{min}^{deviation}$	±5
θ	5
$y_{min}^{deviation}$	±0.2
t_e	36
t_c	126

According to the principle of thrice standard error, the abnormal passenger flow data in the normal historical data can be removed.

C. MODEL PARAMETER CALIBRATION

Combined with historical data and emergency information, emergency data processing parameters and model parameters can be obtained through statistical analysis, as shown in Table 5 and Table 6.

The value of $x_{t,od}^{normal}$ can be sorted from largest to smallest. Make point chart of $x_{t,od}^{normal}$, and the inflection point on the curve of $x_{t,od}^{normal}$ is the parameter θ.

D. EVALUATION OF THE PREDICTION RESULTS

Three indexes (i.e., the mean absolute error (MAE), the root mean square error (RMSE), and the absolute error (AE)) are used to evaluate the effectiveness of the model. The mean absolute percentage error (MAPE) is also commonly used to

TABLE 6. Prediction model parameters.

Parameter	Value	Parameter	Value
λ_1	0.0654	μ_1	39.9267
λ_2	-61.6106	μ_2	-0.0307
λ_3	-92.1902	μ_3	8.0965
λ_4	0.1756	μ_4	-0.0450
λ_5	0.0286	μ_5	1.3138
-	-	μ_6	-39.5156

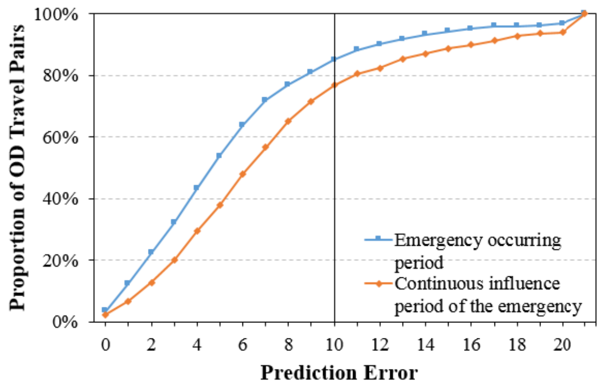


FIGURE 11. OD prediction error distribution.

evaluate the prediction effect, but some OD passenger flow is zero, so the MAPE cannot be calculated in this model.

$$MAE = \frac{1}{n}\sum_{od=1}^{n}\left|x_{t,od}^{emergency} - x_{t,od}\right| \tag{20}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{od=1}^{n}\left(x_{t,od}^{emergency} - x_{t,od}\right)^2} \tag{21}$$

$$AE = x_{t,od}^{emergency} - x_{t,od} \tag{22}$$

At the time granularity of 15 minutes, the MAE values of passenger flow prediction for all the OD pairs in the emergency-occurring period and continuous influence period are 6.62 people and 8.18 people, respectively. The RMSE values of the OD passenger flow are 9.06 people and 9.39 people, respectively, with error deviations of less than 10 people. According to the individual OD pair analyses, the OD pairs with a predicted error of less than 10 people account for a relatively high proportion (85.08%) in the emergency-occurring period and 76.85% in the continuous influence period of the emergency, as shown in Fig. 11. For the 15-minute granularity OD flow data, most of the OD passenger flow prediction errors are between [−10, 10], as shown in Fig. 12 and Fig. 13.

In Fig. 12 and Fig. 13, some OD pairs have a large offset value in passenger flow prediction. By selecting OD pairs with an AE of more than 20 people, we find that 47 and 82 OD pairs have a prediction error of more than 20 people during the emergency-occurring period and continuous influence period, respectively. An analysis of the prediction error in the OD volume exceeding 20 people is shown in Fig. 14. The origin or destination station of the OD trips for which

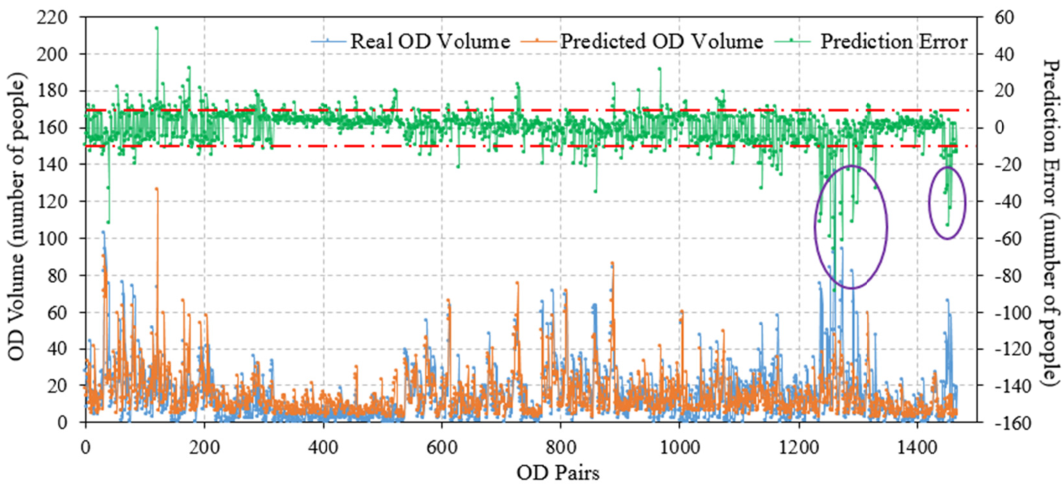


FIGURE 12. Distribution of the predicted values and the true values in the emergency-occurring period.

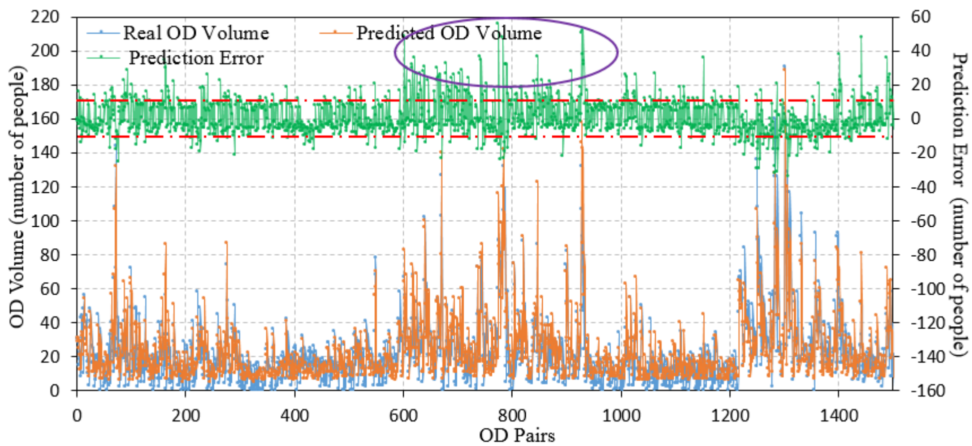


FIGURE 13. Distribution of the predicted values and the true values in the continuous influence period of the emergency.

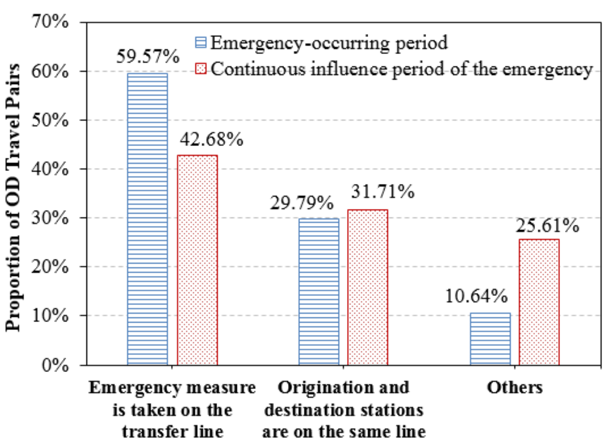


FIGURE 14. OD pair ratio distributions for deviation values exceeding 20 people.

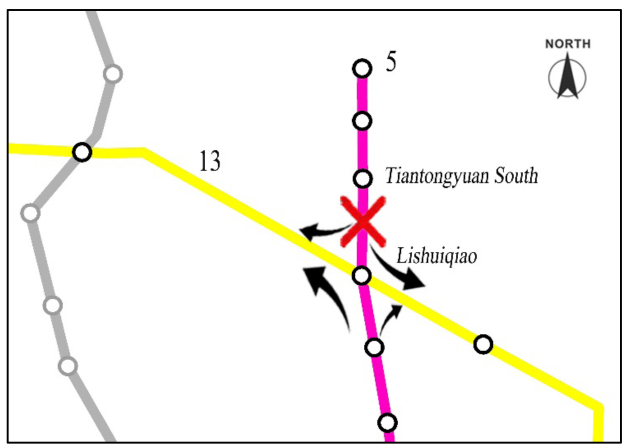


FIGURE 15. Location of the emergency-occurring line and Line 13.

the prediction error exceeds 20 people is mostly on Line 13, accounting for approximately 60% and 40% of both the emergency-occurring period and the continuous influence period, respectively. Lishuiqiao South Station is a transfer station for Line 5 and Line 13. As shown in Fig. 15, there will be much passenger flow from Line 5 to Line 13 after the emergency, but passenger pressure on Line 13 is already quite heavy. To minimize the impact of the emergency on Line 13, the operating enterprise took measures to make the trains of Line 13 skip Lishuiqiao Station. This measure prevented passengers from transferring from Line 5 to Line 13, which severely affected passengers who had planned to travel from Line 5 to Line 13. This study is mainly aimed at the prediction of mass OD passenger flow under the condition that no intervention occurred. For this case, Line 13 becomes the line affected by Line 5, and mass passenger flow on Line 13 can be assumed to be another emergency.

At the same time, 40 OD pairs, which account for approximately 30% of the total OD pairs with an offset error of more than 20 people, are the origin and destination station on the same line in the emergency-occurring period and continuous influence period. In addition, there are 5 and 21 OD pairs with travel prediction errors of more than 20 people in the emergency-occurring period and the continuous influence period, respectively, which have no obvious characteristics. However, compared with the 3169 OD pairs within the influence range, there are 66 OD pairs in these two periods, accounting for approximately 2.1%, which is in the prediction range of allowable error. Therefore, these OD pairs can be ignored.

VI. CONCLUSION AND DISCUSSION

This paper studies the prediction of OD passenger flow in an emergency. The main conclusions are as follows:

(1) Research on emergencies caused by signal faults and vehicle faults during the peak period is the focus of OD passenger flow prediction. Not all trips are affected by the emergency, so it is necessary to determine the influence range in passenger flow prediction research.

(2) The deviation degree of the OD passenger flow in an emergency is strongly correlated with the influence time of emergency, the distance from the origin station or the destination station to the emergency-occurring location, and the shortest distance of OD trip. When modeling the passenger flow prediction, we should start from the two aspects of the emergency-occurring period and the continuous influence period and consider the influence of different occurrence locations, which include a station, a section between two stations or a sections across several stations.

(3) Without considering the influence of the spatial region, the OD passenger deviation rate changing over time in an emergency conforms to a sine function. In the emergency period, the OD passenger flow deviation rate shows a logarithmic distribution with the change in the passenger flow spatial influence parameter, while it shows a power function distribution in the continuous influence period. On this basis, the OD passenger flow prediction model for an emergency is established by using regression analysis with relatively abundant AFC data. The prediction error of the OD passenger flow is less than 10 people at a granularity of 15 minutes.

(4) Due to the nature of land use and line attributes, passenger flow characteristics for different line types are different when an emergency occurs, which results in differences in the parameters of prediction models.

This paper relies on AFC passenger flow data and historical passenger flow rules to predict OD passenger flow under emergencies, which is conducive to effectively induct and evacuate large passenger flow. The proposed method is derived from the regression analysis of passenger flow of BURT based on a large number of AFC data. It can be basically used for URT under the condition of network operation, but the parameters of the model need to be changed according to the actual situation. When verifying the model, OD pairs for which the origin station and destination station are on the same line account for a small proportion with prediction errors of more than 20 people. Therefore, improving the OD passenger flow prediction accuracy for OD pairs with the origin and destination stations on the same line will be one of the focuses in the next step. At the same time, this paper assumes that operating enterprises do not take emergency measures in emergency, which is applicable to most practical situations. However, sometimes, operating enterprises may take appropriate emergency measures to reduce the impact if conditions permit. Therefore, adding the impact of operating enterprises' emergency measures into the model is also one of the directions of model optimization in future research.

REFERENCES

[1] B.-M. Han, W. Dai, and H.-J. Zhang, "Statistics and analysis of urban rail transit operation in the world," *Urban Rapid Rail Transit*, vol. 32, no. 1, pp. 9–14 and 85, Feb. 2019.
[2] W. Hua and G. P. Ong, "Network survivability and recoverability in urban rail transit systems under disruption," *IET Intell. Transp. Syst.*, vol. 11, no. 10, pp. 641–648, Dec. 2017.
[3] S.-S. Liu and E.-J. Yao, "Holiday passenger flow forecasting based on the modified least-square support vector machine for the metro system," *J. Transp. Eng.*, vol. 143, no. 2, Feb. 2017, Art. no. 04016005.
[4] R. Chen, C.-Y. Liang, W.-C. Hong, and D.-X. Gu, "Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm," *Appl. Soft Comput.*, vol. 26, pp. 435–443, Jan. 2015.
[5] H.-D. Yin, B.-M. Han, D. W. Li, and Y. Wang, "Evaluating disruption in rail transit network: A case study of Beijing subway," *Procedia Eng.*, vol. 137, pp. 49–58, Feb. 2016.
[6] L. Hong, J. Gao, and W. Zhu, "Self-evacuation modelling and simulation of passengers in metro stations," *Saf. Sci.*, vol. 110, pp. 127–133, Dec. 2018.
[7] H.-F. Zhao, J. Jiang, R.-Y. Xu, and Y. Ye, "SIRS model of passengers' panic propagation under self-organization circumstance in the subway emergency," *Math. Problems Eng.*, vol. 2014, May 2014, Art. no. 608315.
[8] L. Qu and W. K. Chow, "Platform screen doors on emergency evacuation in underground railway stations," *Tunnelling Underground Space Technol.*, vol. 30, pp. 1–9, Jul. 2012.
[9] Z.-Y. Duan, Z.-X. Lei, M. Zhang, W.-F. Li, J. Fang, and J. Li, "Understanding evacuation and impact of a metro collision on ridership using large-scale mobile phone data," *IET Intell. Transp. Syst.*, vol. 11, no. 8, pp. 511–520, Oct. 2017.
[10] A. Tympakianaki, H. N. Koutsopoulos, E. Jenelius, and M. Cebecauer, "Impact analysis of transport network disruptions using multimodal data: A case study for tunnel closures in Stockholm," *Case Stud. Transp. Policy*, vol. 6, no. 2, pp. 179–189, Jun. 2018.
[11] H.-J. Sun, J.-J. Wu, L.-J. Wu, X.-Y. Yan, and Z.-Y. Gao, "Estimating the influence of common disruptions on urban rail transit networks," *Transp. Res. A, Policy Pract.*, vol. 94, pp. 62–75, Dec. 2016.
[12] L.-M. Zhang, M.-J. Liu, X.-G. Wu, and S. M. AbouRizk, "Simulation-based route planning for pedestrian evacuation in metro stations: A case study," *Autom. Construct.*, vol. 71, pp. 430–442, Nov. 2016.

[13] C.-C. Xu, J. Bao, C. Wang, and P. Liu, "Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China," *J. Saf. Res.*, vol. 67, pp. 65–75, Dec. 2018.

[14] X.-G. Chen and H.-P. Deng, "A correlation analysis of information use, social networks and cooperation consciousness in travel behaviors," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 62, pp. 819–832, Apr. 2019.

[15] S. Das and X.-D. Sun, "Association knowledge for fatal run-off-road crashes by multiple correspondence analysis," *IATSS Res.*, vol. 39, no. 2, pp. 146–155, Mar. 2016.

[16] T. L. Pan, A. Sumalee, R.-X. Zhong, and N. Indra-Payoong, "Short-term traffic state prediction based on temporal–spatial correlation," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1242–1254, Sep. 2013.

[17] Y. Wang, Y. Zhang, X.-L. Piao, H. Liu, and K. Zhang, "Traffic data reconstruction via adaptive spatial-temporal correlations," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 4, pp. 1531–1543, Aug. 2018.

[18] X. Wang, Q. Li, J.-F. Yuan, and P. M. Schonfeld, "Metro passenger behaviors and their relations to metro incident involvement," *Accident Anal. Prevention*, vol. 82, pp. 90–100, Sep. 2015.

[19] A. J. Horowitz, "Simplifications for single-route transit-ridership forecasting models," *Transportation*, vol. 12, pp. 261–275, May 1984.

[20] X.-L. Ma, J.-Y. Zhang, B.-W. Du, C. Ding, and L.-L. Sun, "Parallel architecture of convolutional bi-directional LSTM neural networks for network-wide metro ridership prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2278–2288, Jun. 2019.

[21] Y. Liu, Z. Liu, and R. Jia, "DeepPF: A deep learning based architecture for metro passenger flow prediction," *Transp. Res. C, Emerg. Technol.*, vol. 101, pp. 18–34, Apr. 2019.

[22] Y. Li, X.-D. Wang, S. Sun, X.-L. Ma, and G.-Q. Lu, "Forecasting short-term subway passenger flow under special events scenarios using multi-scale radial basis function networks," *Transp. Res. C, Emerg. Technol.*, vol. 77, pp. 306–328, Apr. 2017.

[23] M. Ni, Q. He, and J. Gao, "Forecasting the subway passenger flow under event occurrences with social media," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 6, pp. 1623–1632, Jun. 2017.

[24] J.-Y. Guo, Z. Xie, Y. Qin, L.-M. Jia, and Y.-G. Wang, "Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM," *IEEE Access*, vol. 7, pp. 42946–42955, Mar. 2019.

[25] J.-Y. An, L. Fu, M. Hu, W.-H. Chen, and J.-W. Zhan, "A novel fuzzy-based convolutional neural network method to traffic flow prediction with uncertain traffic accident information," *IEEE Access*, vol. 7, pp. 20708–20722, Feb. 2019.

[26] S.-S. Liu, E.-J. Yao, B.-B. Li, and Y. Tang, "Forecasting passenger flow distribution between urban rail transit stations based on behavior analysis under emergency events," *J. China Railway Soc.*, vol. 40, no. 9, pp. 22–29, Sep. 2018.

[27] B. Wang, *Theory and Application for Urban Rail Transit Network Operation*. Beijing, China: Commun. Press Co., Ltd, 2014, p. 60.

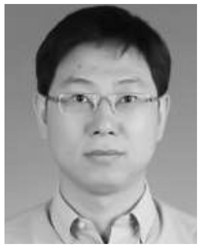
[28] J. Louie, A. Shalaby, and K. N. Habib, "Modelling the impact of causal and non-causal factors on disruption duration for Toronto's subway system: An exploratory investigation using hazard modelling," *Accident Anal. Prevention*, vol. 98, pp. 232–240, Jan. 2017.

[29] J.-H. Wang, W.-Y. Yan, Y.-R. Zhi, and J.-C. Jiang, "Investigation of the panic psychology and behaviors of evacuation crowds in subway emergencies," *Procedia Eng.*, vol. 135, pp. 128–137, Feb. 2016.

[30] L.-Y. Tang, Y. Zhao, J. Cabrera, J. Ma, and K. L. Tsui, "Forecasting short-term passenger flow: An empirical study on shenzhen metro," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3613–3622, Oct. 2019.

CHEN LI received the M.S. degree from the Beijing University of Technology, Beijing, China, in 2017, where he is currently pursuing Ph.D. degree in transportation engineering with the College of Metropolitan Transportation. His research interests include rail transit safety, travel behavior analysis, and optimization modeling.

JIANLING HUANG received the Ph.D. degree from Beijing Jiaotong University, Beijing, China, in 2011. She is currently a Professor of engineering, the Director of the Beijing Transportation Information Center, and the Ph.D. Supervisor with the Beijing University of Technology. Her research interests include transportation information, big data, and intelligent transportation.



BO WANG received the Ph.D. from Beijing Jiaotong University, Beijing, China, in 2008. He is currently a Professor of engineering with the Beijing Transportation Information Center and the Ph.D. Supervisor with Beijing Jiaotong University. His research interests include the organization of train operations, data mining, and passenger flow prediction.

YUYANG ZHOU received the Ph.D. degree in operational research and cybernetics from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2011. She is currently an Associate Professor and the Deputy Head of the Department of Transportation Engineering, Beijing University of Technology, Beijing. Her research interests include intelligent transport system problems jointed with the (big) data analysis.

YUNYUN BAI received the M.S. degree from Beijing Jiaotong University, Beijing, China, in 2015. She is currently a Senior Engineer with the Beijing Transportation Information Center. Her research interests include the organization of train operations, data mining, and passenger flow prediction and analysis for urban rail transit.

YANYAN CHEN received the Ph.D. degree from the Harbin Institute of Technology, Harbin, China, in 1997. She is currently a Professor of engineering, the Dean of the College of Metropolitan Transportation, and the Ph.D. Supervisor with the Beijing University of Technology, Beijing, China. Her research interests include transportation planning, big data, and intelligent transportation.