

Article

The Predictability of Short-Term Urban Rail Demand: Choice of Time Resolution and Methodology

Zi-jia Wang ¹, Hai-xu Liu ¹, Shi Qiu ^{2,*}, Ji-ping Fang ³ and Ting Wang ⁴¹ School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China; zjwang@bjtu.edu.cn (Z.-j.W.); haixuliu@bjtu.edu.cn (H.-x.L.)² School of Civil Engineering, Central South University, No. 22 Shaoshan South Road, Changsha 410075, China³ Centre for Transport Studies, Department of Civil, Environmental and Geomatic Engineering, University College London, London WC1E6BT, UK; jiping.fang.15@ucl.ac.uk⁴ China Railway First Survey and Design Institute Group Co., Ltd., Xi'an 710073, China; 15121189@bjtu.edu.cn

* Correspondence: sheldon.qiu@bjtu.edu.cn

Received: 30 September 2019; Accepted: 31 October 2019; Published: 5 November 2019



Abstract: The accuracy of short-term demand forecasting is critical for real-time operation management of urban rail transit, which largely depends on the choice of time resolution. Although there have been continuous improvements in forecasting models, the basic issue has not been well addressed. In this regard, the predictability of short-term demand in terms of time resolution setting and the corresponding model selection have been addressed in this study. Two methods have been considered: the demand forecasting with the past demand during the same time slot on the same weekday (the same period method); and that with continuous time series demand exactly before the forecasted time slot (time series method). The predictability for these two methods was respectively measured by the similarity of the same period and the stability of the time series. Consequently, the influence of time resolution on the predictability of short-term demand for urban rail transit has been evaluated. With the methods proposed, this study conducted an analysis on five-week smartcard data in the Beijing subway system. Results suggest that the predictability of short-term demand presented remarkable heterogeneity in both time and space. The predictability of demand forecasting at station level has been summarized into different levels, and the corresponding methods can be selected for each class. Generally, to ensure a desirable accuracy, forecasting can be made at a 10-min and 60-min interval on weekdays and weekends, respectively. The same period method works better for the short-term demand forecasting on weekdays. While the time series method performs better for prediction on weekends. As for short-time OD (origin-destination) demand, the time series method with a 10-min interval, which is supplemented by the same period method, can generate acceptable forecasting results. In brief, this study provides suggestions on the time resolution and method selection for short-term demand forecasting.

Keywords: short-term demand; time series; time resolution; similarity measurement; stability test

1. Introduction

In recent years, urban rail transit in China has developed rapidly, leading to network expansion. Many urban rail transit networks have been challenged by issues in the management of passenger demand [1]. It requires a more accurate model of short-term prediction for urban rail transit network operation. According to the difference in prior passenger information, the short-term demand prediction methods can be divided into the same period method and the time series method. The main prediction objects are station boarding demand and OD (Origin-Destination) demand. Since choosing time resolution for such kind of forecasting is a basic issue in short-term demand prediction, it affects the prediction accuracy directly.

Various methods have been used in the field of prediction for short-term transport demand and OD demand, though not limited to rail transit [2,3]. According to the two type of data papers, time series data and the same period data, we can divide literatures into two categories, times series method and the same period method.

1.1. Time Series Method

Based on the time series data of passenger flow, lots of models have been used in short-term predication. In particular, the autoregressive integrated moving average (ARIMA) is mostly used. It assumes the current demand is a linear combination of some orders of demand in previous time slots. To take the apparent seasonal features of transport demand into account, SARIMA (seasonal autoregressive integrated moving average) was established to achieve better results [4]. Such type of linear model cannot capture the non-linear features of transport demand. Thus, the non-linear model emerged, such as neural network and support vector machine models (SVM) [5,6].

In recent years, using combined models rather than a single model became a trend in this field [7]. Zhang et al used a series of station boarding demand at 2-min interval in the morning-peak (6:00–9:05) of 5 weekdays to predict the ridership [8]. These combined models showed better performance compared to single models. Sun et al developed a wavelet-SVM model to predict demand and interchange demand based on 15-min interval [9]. It was suggested that the wavelet-SVM model was the most suitable short-term demand forecasting model for the Beijing subway system and had strong robustness. Wei et al proposed a hybrid model of empirical mode decomposition and neural network to predict short-term metro demand using time series demand records [10]. The three-stage model uses the empirical mode decomposition model to extract inputs from the original dataset at the first two stages, and the prediction is made with these inputs fed to a neural network model at the final stage. The proposed model outperforms the existing model such as SARIMA. Jia et al also formulated a combined model with GM (grey model) and ARMA (autoregressive moving average), using the time series data of passenger flow in the weekdays in four continuous weeks to forecast the demand at a subway station every 10 min [11].

1.2. The Same Period Method

One of the shortages of the time series method is that it cannot realize the different type days, such as weekdays and weekends or festivals and ordinary days, which causes the models based on time series data to not show good performance all the time. So, the same period data has been used to solve the problem. Yao et al forecasted the OD demand in the intervals of 15 min, 30 min, and 60 min with data extracted from the AFC (auto fare collection) data of the same period last week [12]. The result showed that prior information credibility increased with the increase of the interval length, as well as its accuracy, especially during the peak hours. The interval should be 15 min in order to meet the requirements of estimated intervals and estimation precision. Du et al analyzed the fluctuation and similarity of expressway OD under multi-intervals [13]. The results showed that it was difficult to accurately predict transit demand within 2 h-intervals using the time series demand in last slots. If using the same period historical data to forecast the OD demand within the scope of the expressway, choosing time resolution of 30 min would yield a higher prediction accuracy, especially on weekdays.

The prediction models and intervals of short-term passenger flow differed greatly in the previous studies. But they have ignored two basic questions in short-term prediction. One is whether the short-term demand of each station is predictable. What is a proper predicting method for a station which has large instability in the prior passenger demand? To what extent does the demand show a strong regularity? The other question is how to select the time resolution to make the models show better performance? Shall we consider different intervals to predict demand in different time periods (peak and off-peak)? Zhong analyzed the pattern of passenger flow and OD demand with different time resolutions based on one-week AFC data of London, Singapore, and Beijing. The results indicate that the minimum interval for forecasting meeting the requirement of accuracy should be 15 min. Larger

time resolutions will sacrifice the predicting accuracy to various degrees [14]. That work indicates there are limitations on time resolution selection.

To address these basic questions, this study has evaluated the predictability of different forecasting objects in the entire network scale and analyzed the choice of intervals and methods for short-term demand forecasting. The remaining part of this article is structured as follows. The next section describes the dataset used and the methodologies established in this work. Then, the results are presented for both entry demand and ODs. These results are also visualized in GIS (geographic information system), followed by the discussions on the results in terms of forecasting method selection and time resolution selection. Finally, conclusions and suggestions on rail transit short-term demand forecasting are provided.

2. Methodology

In this section, firstly we give a brief introduction of the dataset. Then, the methods applied in this study are formulated in details. The purpose of this study is to investigate the method and time resolution choice problem of short-term rail ridership forecasting. According to the prior data used for demand forecasting in the next time slot, the philosophy of the predicting method can be roughly classified into two categories: ‘same period’ method and ‘time series’ method. Hereafter the ‘same period’ method forecasts urban rail demand according to the ridership in the same time slot of day in a continuous few days or in the same time slot on the same weekday of several continuous weeks. The ‘time series’ method refers to forecasting passenger demands based on the historical demand data in a few time resolutions exactly before the target time slot. For the first method, similarity in terms of PCC (Pearson correlation coefficient) is normally used to measure the predictability; while for the latter one, time stability is commonly used.

2.1. Dataset

The dataset used in the study was the subway AFC data of 5 consecutive weeks from March 2016 to April 2016. After removing redundant data, the dataset structure is shown in Table 1. Card_id is a card number that represents the unique identity of a passenger. Entry_station and Exit_station are the boarding stations and the exiting stations of the passenger, respectively. Entry_time and Exit_time ranged from 1 to 1140 are the time series numbers of entering station and exiting station respectively, because the metro operating hours are 5:00–24:00 every day, 1140 min a day in total. Then, the time series of station boarding passenger flow and OD passenger flow are extracted.

Table 1. An example of AFC (Auto Fare Collection) data after cleaning on a certain day in March.

Card Id	Entry Station	Exit Station	Entry Time	Exit Time
190604	260	9	397	439
128477	19	22	826	840
390467	15	244	704	734
197686	191	115	279	314
752308	143	247	193	229

2.2. Same Period Similarity Measurement

According to [14], 15 representative intervals are selected as $\Delta t(\text{min})$:

$$\Delta t = [1, 2, 5, 10, 15, 20, 30, 60, 95, 114, 190, 228, 285, 380, 570].$$

Then, 1140 min of a day (5:00–24:00) are divided into 15 intervals:

$$n = [1140, 570, 228, 114, 76, 57, 38, 19, 12, 10, 6, 5, 4, 3, 2].$$

The vector $X_{N_Δt}(i_D)$ is used to represent time series of station boarding flow:

$$X_{N_Δt}(i_D) = [x_1, x_2, x_3, \dots, x_t, \dots, x_n],$$

where N denotes the station number; $Δt$ denotes interval; i_D denotes the day D of week i ; $i \in (1, 5)$ denotes that the AFC data is selected for five consecutive weeks; $D \in (1, 3, 5, 7)$ denotes selecting Monday, Wednesday, Friday, and Sunday from each week; $t \in (1, n)$ denotes 1140 min, per day is divided into n segments $n = 1140/Δt$.

The PCC [15] is used to measure the similarity during the same period of two weeks. For example, the Pearson coefficient index $r_{N_Δt}(i_D, j_D)$ of $x_{N_Δt}(i_D)$ and $x_{N_Δt}(j_D)$, which is the coefficient of passenger flow time series on the same day D of the week i and week j under the interval $Δt$, can be calculated as following:

$$r_{N_Δt}(i_D, j_D) = \frac{\sum_{k=1}^n (x_{N_Δt_k}(i_D) - \overline{x_{N_Δt}(i_D)})(x_{N_Δt_k}(j_D) - \overline{x_{N_Δt}(j_D)})}{\left(\sqrt{\sum_{k=1}^n (x_{N_Δt_k}(i_D) - \overline{x_{N_Δt}(i_D)})^2}\right)\left(\sqrt{\sum_{k=1}^n (x_{N_Δt_k}(j_D) - \overline{x_{N_Δt}(j_D)})^2}\right)}.$$

Because of $i, j \in (1, 5)$, measuring the similarity between the same day D of five weeks yields a 5×5 symmetric matrix consisting of 10 Pearson coefficients:

$$R_{N_Δt}(i_D, j_D) = \begin{bmatrix} 1 & r_{N_Δt}(2_D, 1_D) & r_{N_Δt}(3_D, 1_D) & r_{N_Δt}(4_D, 1_D) & r_{N_Δt}(5_D, 1_D) \\ r_{N_Δt}(1_D, 2_D) & 1 & r_{N_Δt}(3_D, 2_D) & r_{N_Δt}(4_D, 2_D) & r_{N_Δt}(5_D, 2_D) \\ r_{N_Δt}(1_D, 3_D) & r_{N_Δt}(2_D, 3_D) & 1 & r_{N_Δt}(4_D, 3_D) & r_{N_Δt}(5_D, 3_D) \\ r_{N_Δt}(1_D, 4_D) & r_{N_Δt}(2_D, 4_D) & r_{N_Δt}(3_D, 4_D) & 1 & r_{N_Δt}(5_D, 4_D) \\ r_{N_Δt}(1_D, 5_D) & r_{N_Δt}(2_D, 5_D) & r_{N_Δt}(3_D, 5_D) & r_{N_Δt}(4_D, 5_D) & 1 \end{bmatrix}.$$

In the comparison, the average of 10 Pearson coefficients is used as a similarity indicator for station boarding passenger flow:

$$C_{N_D_Δt} = \sum_{i=1}^l \sum_{j=i+1}^l (r_{N_Δt}^2(i_D, j_D) / (l(l-1)/2)),$$

where $r_{N_Δt}^2(i_D, j_D)$ represents similarity between $x_{N_Δt}(i_D)$ and $x_{N_Δt}(j_D)$. For instance, $r_{210_5}^2(2_3, 5_3) = 0.7$ denotes that the station boarding passenger flow (for station 210) on Wednesday in the second week has a similarity of 70% with the same day in the fifth week when the interval is 5 min. The closer the degree of similarity is to 1, the more significant the passenger flow pattern is. l is the number of measured weeks, in this study $l = 5$. The measured similarity of the entry passenger volumes at all the stations in the entire network for every day D in measured weeks is a one-dimensional matrix:

$$RR_{D_Δt} = [C_{1_D_Δt}, C_{2_D_Δt}, \dots, C_{N_D_Δt}].$$

Variables in Section 2.2 are shown in Table 2.

Table 2. Variables in Section 2.2.

Variables	Type	Meanings
Δt	Number	Time resolution
n	Number	The number of intervals a day
l	Number	the number of measured weeks
$X_{N_ \Delta t}(i_D)$	Vector	time series of station N boarding with the time resolution of Δt in the day D of week i
$r_{N_ \Delta t}(i_D, j_D)$	Number	the coefficient of passenger flow time series on the same day D of the week i and week j under the interval Δt
$R_{N_ \Delta t}(i_D, j_D)$	Matrix	a 5×5 symmetric matrix consisting of 10 Pearson coefficients
$r_{N_ \Delta t}^2(i_D, j_D)$	Number	similarity between $X_{N_ \Delta t}(i_D)$ and $X_{N_ \Delta t}(j_D)$
$C_{N_ D_ \Delta t}$	Number	the average of 10 Pearson coefficients
$RR_{D_ \Delta t}$	Vector	The measured similarity of the entry passenger volumes at all the stations in the entire network for every day D in measured weeks

3. Time Series Stability Test

The augmented Dickey–Fuller test (ADF) is commonly used to determine if time series data is stable or not [16]. A stable time series is easier to predict with the very complex model [17]. Thus, this index can also be used as an indicator of predictability. The process of using the ADF test to evaluate the stability of the passenger flow time series is shown in Figure 1. There are three forms in the ADF test regression model: constant term and time trend term included, only constant term included and no constant term or time trend term. Each has its own threshold value, and an appropriate regression model can be selected based on the curve of the sequence. The equation of the general time series is shown below:

$$Y_t = \alpha_1 + \alpha_2 t + \sum_{i=1}^m \beta_i Y_{t-i} + \varepsilon_t.$$

In the equation, t is the time variable; m is the number of intervals; Y_t refers to the demand at time slot t ; α_1 is a constant term; $\alpha_2 t$ denotes a time trend term; β_i is i th coefficient; Y_{t-i} refers to the demand at time slot $t-i$; ε_t is a residual.

From the equation above we can derive the 3 forms of the equations as follows:

(1) Constant term and time trend term included:

$$\Delta Y_t = \alpha_1 + \alpha_2 t + \rho Y_{t-1} + \sum_{i=1}^m \beta_i \Delta Y_{t-i} + \varepsilon_t.$$

(2) Only constant term included:

$$\Delta Y_t = \alpha_1 + \rho Y_{t-1} + \sum_{i=1}^m \beta_i \Delta Y_{t-i} + \varepsilon_t.$$

(3) No constant term or time trend term:

$$\Delta Y_t = \rho Y_{t-1} + \sum_{i=1}^m \beta_i \Delta Y_{t-i} + \varepsilon_t.$$

In the above three equations, ρ is a parameter.

In the inspection process, the sequence diagram is firstly observed. If the time trend term and constant term exist, select the first form; if the constant term is significant, select the second form; if neither the time trend nor the constant is significant, select third form with no constant term and time trend.

Using the t-test with the Hypothesis $H_0 : \rho = 0$, the resulting value is the ADF value P . Given the significance level, if the ADF statistic value P is less than the threshold value, the parameter ρ is significantly different from 0, and unit root does not exist in the sequence Y_t , which means the sequence is stable.

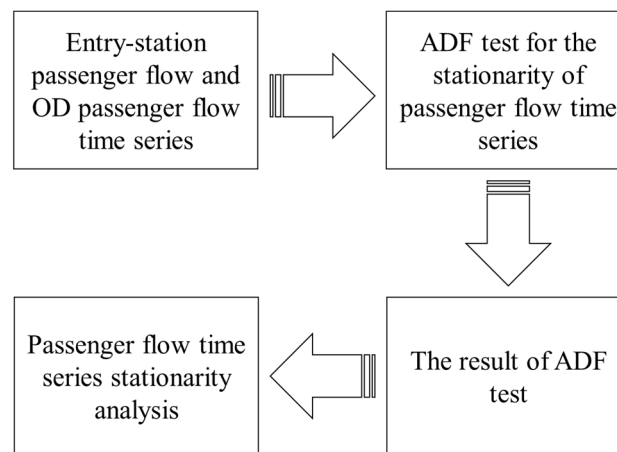


Figure 1. Passenger flow time series stability test process.

In order to ensure the accuracy of different ADF stability evaluation test results, the corresponding estimated value P of ADF is used as the evaluation index. Under the significance level of $\alpha = 0.05$, if $P \leq \alpha$, the original hypothesis is rejected and the sequence does not have a unit root, which means the sequence is significantly stable. The less the P value, the more significant the stability of the sequence; if $P > \alpha$, the original hypothesis is accepted, the sequence has a unit root, which suggests the sequence is not stable, and the larger the P value, the less significant the stability of the sequence.

Selecting one of the weekdays and one of the weekends in a week as the characteristic days D in the study, $n = 1140/\Delta t$, then there is a P value for each station passenger flow (OD) time series $X_n(i_D)$ after testing based on each intervals Δt . Therefore, the passenger flow (OD) stability test result at each station in the entire network on the day D of a week is a one-dimensional matrix:

$$P_{D_{\Delta t}} = [P_1, P_2, P_3, \dots, P_N].$$

4. Experimental Results

4.1. Entry Passenger Volumes at Station Level

4.1.1. Similarity Measure between Entry Passenger Volumes of the Same Period

The higher the similarity of entry passenger volumes in the same day of different weeks, the higher the predictability of short-term entry demand. Therefore, it is important to explore the minimum time resolution for station short-term demand forecasting. In Figure 2, y axis means the percentage of stations which meet the requirements in total stations, x axis means different time resolution, $RR_{D_{\Delta t}} > 0.75$ and $RR_{D_{\Delta t}} > 0.9$ show the percentage of stations with similarity of $C_{N_{D_{\Delta t}}} > 0.75$ and $C_{N_{D_{\Delta t}}} > 0.9$ in the total number of stations in different time resolutions Δt . It can be found that the proportion of stations with similarity greater than 0.75 and 0.9 is increasing with the interval over time. To ensure that 90% of the stations in the entire network have a high similarity ($RR_{D_{\Delta t}} > 0.9$), the minimum time resolutions which ensure the accuracy of demand forecasting are 15 min on weekdays, and 60 min on weekends. To ensure that 90% of the stations in the entire network have a strong similarity ($RR_{D_{\Delta t}} > 0.75$), the minimum time resolutions should be 5 min on weekdays and 20 min on weekends. Overall, the larger the time resolution, the stronger the similarity of the station's short-term station boarding demand. By studying the variation pattern of similarity of station boarding demand with time, the "same period" method should be used to predict the optimal time resolution of short-term entry-demand at a station.

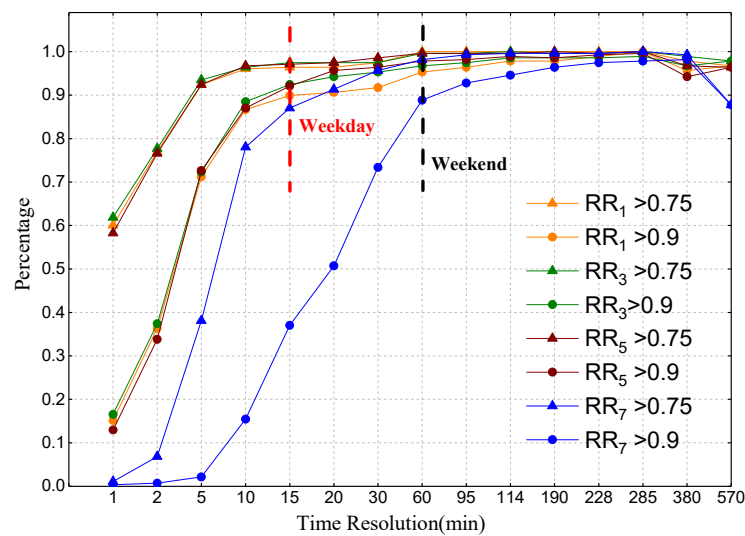


Figure 2. Comparison between the predictabilities of the entry passenger volumes in different time resolutions using the same period method.

4.1.2. Time Series Stability Test of Entry Passenger Volumes

Figure 3 shows the change of the proportion of stable stations with time resolution at three significant levels in weekdays and weekends. It can be seen that the proportion of stations with stable demand rises quickly in 10 min and reached a plateau afterwards within 60 min on weekdays. At 10 min time resolution, the proportion of stations with stable demand reach peak and stay stable, at this point, the proportion under three kinds of significance test, $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.1$, which are widely used in statistics, are 39%, 62%, and 75%, respectively, indicating that 62% of the station traffic can get accurate prediction results by the “time series” method based on 10-min time resolution. During the weekends, the proportion of stations with stable demand decreases with the increase of time resolution within 30 min. The proportion remains growing from 30 min to 114 min, until reaching the maximum value. When the time resolution is 114 min and the α has the same values, the proportion of station with stable demand are 15%, 30%, and 43%, respectively under three significance tests. Not only is the stable demand station proportion far lower than the weekdays, but also the time resolution is far larger than the one for weekdays, which is 10 min. Therefore, it is difficult to obtain accurate prediction results even at large time resolution using the “time series” method for predicting short-time demand on weekends.

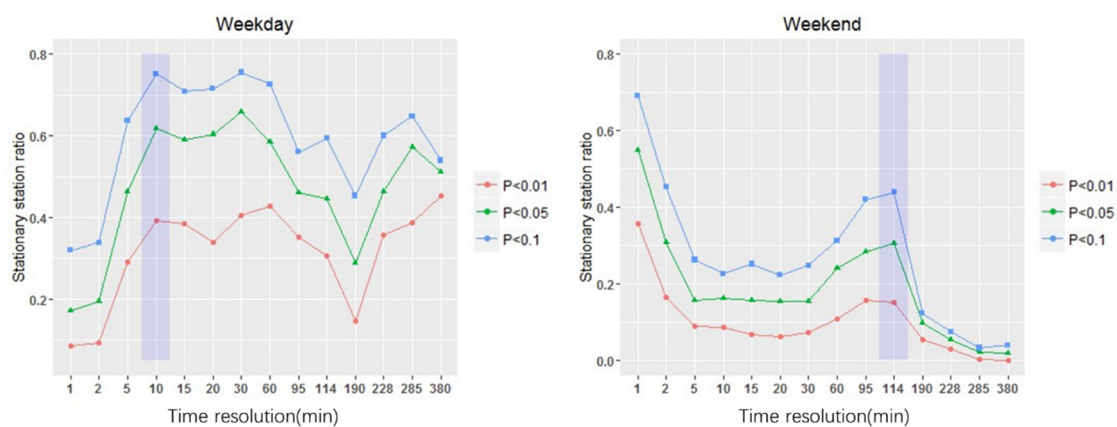


Figure 3. The proportion of steady station boarding passenger flow at different levels on weekday and weekend.

The trend of passenger flow on weekends is obviously different from that on weekdays. Many stations have a stable demand on small time resolution (1 min, 2 min) on weekends, which is related to fewer passengers on weekends. It is also related to the strong randomness of travel time and place. Usually it takes a few minutes to complete uploading the smart card information regarding the passenger flow. Therefore, the short time interval has a low operability in real time in short-term prediction. On average, using the “time series” method to forecast short-term demand on weekends can only make accurate predictions for 24% of the stations (about 67 stations) with 60-min interval. Although increasing the interval to 114 min, which is too large, would make the proportion of stations with stable demand increase to 30%, it loses the meaning of short-term demand prediction.

4.2. Passenger Volumes in Different OD Pairs

Figure 4 shows the distribution of OD demand percentiles in the four selected days of a week. The OD pairs with passenger volume over 0 exceed 72,000. On weekdays, the maximum passenger volume exceeds 6000, and on weekends, the maximum passenger flow is about 2000, but the passenger volume in 80% daily OD is less than 100. Since the minimum time resolution chosen in the study is 1 min, it is of little significance to study the OD pairs with low passenger volumes. Therefore, the top 200 passenger volumes ODs in weekdays are studied. The similarity and the stability for the passenger volumes in these 5000 OD pairs are measured and tested.

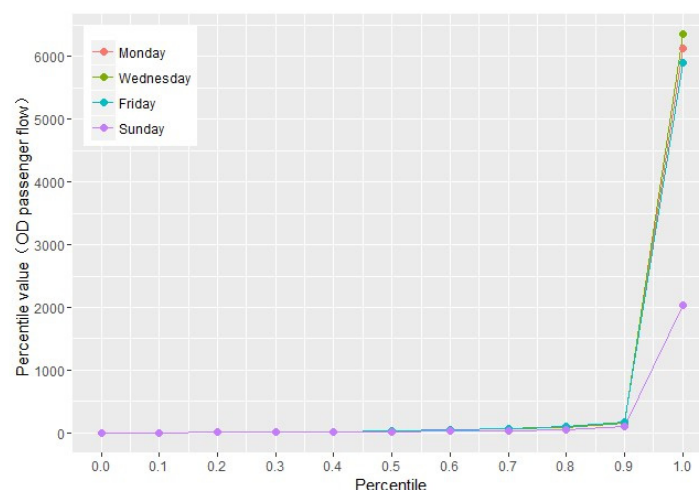


Figure 4. OD (Origin-Destination) passenger flow percentile distribution for different weeks.

4.2.1. Similarity Measure of the Same Period

In Figure 5, $RR_{D_Δt}$ of 0.75 and 0.90 are also used as the standard for division, and the proportion of OD pairs with strong predictability and high predictability is calculated. As can be seen from the figure, the larger the time resolution, the more OD pairs enter the range where predictability is strong and high. From the growth rate of the curve, it can be seen that on weekdays, the growth rate decreases at the 60-min interval. At this interval, the proportion of ODs with strong predictability ($RR_{D_Δt} > 0.75$) exceeds 90%, but the proportion of OD pairs with high predictability ($RR_{D_Δt} > 0.9$) is low, about 65%; at the 190-min interval on weekends, the growth rate reduces, the proportion of ODs with strong predictability exceeds 60%, and the proportion of ODs with high predictability is only higher than 20%. Thus, the predictability of OD demand is low on both weekdays and weekends, and the OD demand on weekends cannot be accurately predicted with a small interval. However, on weekdays, it is possible to make accurate predictions for OD demand with 60-min interval. Compared with the analogy of station boarding demand, the OD demand is more difficult to predict. The most important reason is that the OD demand forecast not only pays attention to where the passengers come from, but also to which stations the passengers are travelling to, which increases the uncertainty. Another

reason is that the number of OD pairs is much higher than the number of stations, and the passenger volume of a single OD is much smaller than that of a single station. Overall, on weekends, it is not suitable to predict the short-term OD demand with the same period method, while on weekdays, it is more reasonable to set the interval at 60 min to predict the short-term OD demand with the same period method.

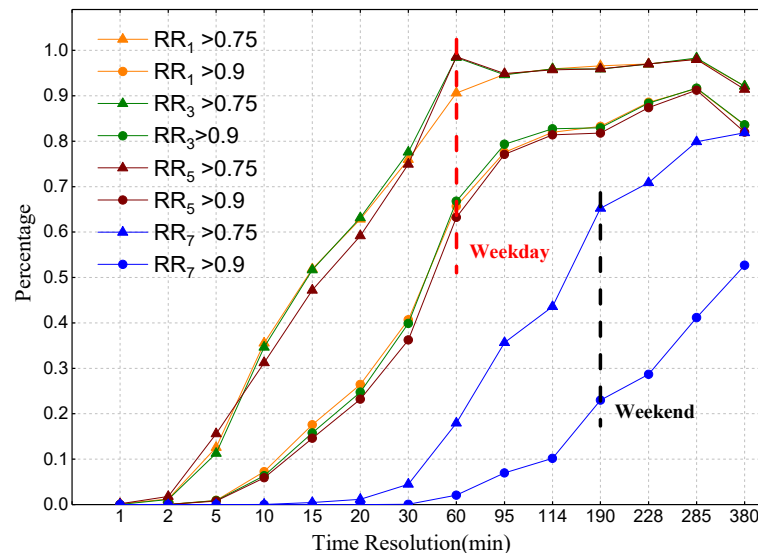


Figure 5. Comparison and analysis of predictability of passenger volume in different OD in same period.

4.2.2. Time Series Stability Test

The relationship between the proportion of stable OD demand and the time resolution of the OD demand on weekdays and weekends under the three levels of significance test is shown in Figure 6. It can be seen that on weekdays and weekends, the proportion of ODs with stable demand decreases when interval increases. Therefore, the larger the time resolution, the larger the variability of OD demand. And the downward trend on weekdays is larger than that on weekends. Taking into account the feasibility of time resolution in actual forecasting, it is reasonable to choose 10-min time resolution on weekdays and weekends. Meanwhile, the corresponding proportions of ODs with stable flow at the three levels of significance test where $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.1$ are 29.38%, 57.38%, and 77.76%, respectively on weekdays, and 81.86%, 87.68%, and 90.80% on weekends. Thus, the proportion of ODs with stable flow on weekends is significantly higher than that on weekdays by 30.30%.

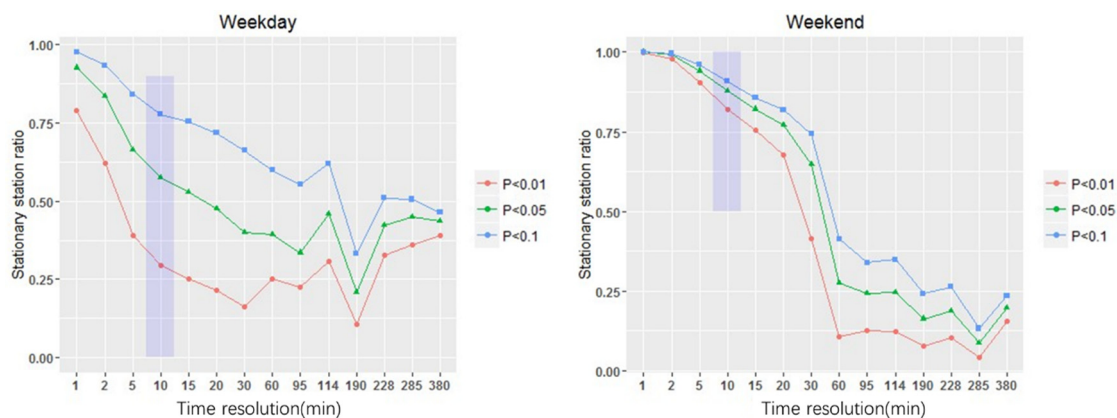


Figure 6. The proportion of steady OD passenger flow at different levels on weekday and weekend.

5. Discussion

The higher the similarity of passenger flow time series and the smaller the variability, the more reliable the results are. Based on this principle, once the thresholds of similarity and stability are set, the predictable ranks of short-time station boarding and OD passenger flow are determined.

5.1. Forecasting Boarding Demand at Station Level

5.1.1. Comparison of Metrics for Weekdays and Weekends

Figure 7 compares the similarity measurement and stability test results of the station boarding passenger flow and shows a scatter plot for stations with different attributes. Ten-min interval on weekday and 60-min interval on weekend are selected to compare the same period method with the time series method for predicting applicability. The red-shaded area indicates that the station boarding passenger flow has a larger similarity ($RR_{D_\Delta t} > 0.9$), which is suitable for the same period forecasting method. The blue-shaded area indicates that the station boarding passenger flow has smaller variability ($P < 0.05$), and is suitable for using time series method for prediction. Overlapping regions indicate that both methods are applicable, while other regions indicate that using either of the two methods alone cannot achieve accurate predictions.

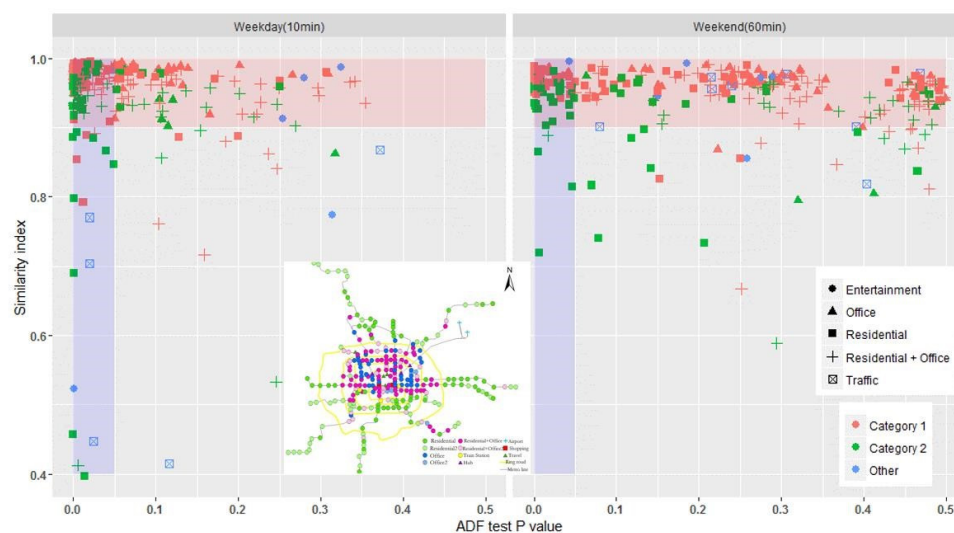


Figure 7. Comparison between the same period and time series predictions of station boarding passenger flow on weekday and weekend.

With the comparison between entry passenger volumes on weekdays and weekends, it can be seen that on weekdays, the stations are mostly clustered in the upper left corner in Figure 7, with only a few stations outside of the shaded area, and on weekends, stations with greater variability have a larger similarity. Taking the attributes of stations into account, on weekdays, stations with the attributes of “residential” and “office” have a relatively high stability and the same period similarity, which is related to the commuting characteristics of such stations. “Traffic” attribute includes “train station”, “hub”, and “airport”. On weekdays, stations with this attribute have a lower same period similarity and a less fluctuant passenger flow. While on weekends, stations with this attribute have a higher same period similarity and a more fluctuant passenger flow. This is related to the passenger flow carried by stations with “traffic” attributes having different nature on weekdays and weekends. “Entertainment” attribute includes “shopping” and “travelling”. Stations with this attribute are only a few and have a high same period similarity.

5.1.2. Predictability Level for Boarding Demand

As is shown in Table 3, based on the same period similarity and the ADF stability of the station boarding passenger time series, the predictability of short-term station boarding passenger flow at the station is ranked.

Table 3. The level of the predictability of short-term station boarding passenger flow at the station.

Level	Feature
Rank 1	Stations that are suitable for both methods. The same period similarity $RR_{D_Δt} > 0.9$ and the time series $P < 0.05$ for the ADF test. These stations are in the overlapping area of the blue and red shade in Figure 7.
Rank 2	Stations that are suitable for the same period method. The same period similarity $RR_{D_Δt} > 0.9$ and the time series $P > 0.05$ for the ADF test. These stations are in the area that is covered by red shade alone in Figure 7.
Rank 3	Stations that are suitable for the time series method. The same period similarity $RR_{D_Δt} < 0.9$ and the time series $P < 0.05$ for the ADF test. These stations are in the area that is covered by blue shade alone in Figure 7.
Rank 4	Stations that are not suitable for either of the methods. The same period similarity $RR_{D_Δt} < 0.9$ and the time series $P > 0.05$ for the ADF test. These stations are in the unshaded area in Figure 7.

Station Rank Changes on Different Days of the Week

Whether the same station has the same predictable rank on different days of the week is important in understanding the passenger flow pattern of stations. Figure 8 shows the change of the predictable rank of the short-term station boarding passenger flow among different days of the week. On weekdays, more stations change predictable ranks from Rank 1 (on Monday) to Rank 2 (on Wednesday) and from Rank 2 (on Wednesday) back to Rank 1 (on Friday), which indicates that the passenger flow on Wednesday is more volatile than that on Monday and Friday. Considering the rank stability of stations, 56.4% of the stations remained unchanged during the three weekdays, of which stations in Rank 1, Rank 2, Rank 3, and Rank 4 accounted for 39.9%, 13.3%, 1.4%, and 1.8%, respectively in these three weekdays. Generally, the predictable ranks of the three feature days on weekdays fluctuated little, and there are more rank changes from weekdays to weekends. For example, more than half of the stations belonging to Rank 1 on Friday fell to Rank 2 on Sundays. Even though the time resolution of weekends is 60 min, the station boarding passenger flow fluctuations are still large. Therefore, in order to obtain better prediction results, it is necessary to select appropriate time resolution and forecasting method for stations of different days and different ranks.

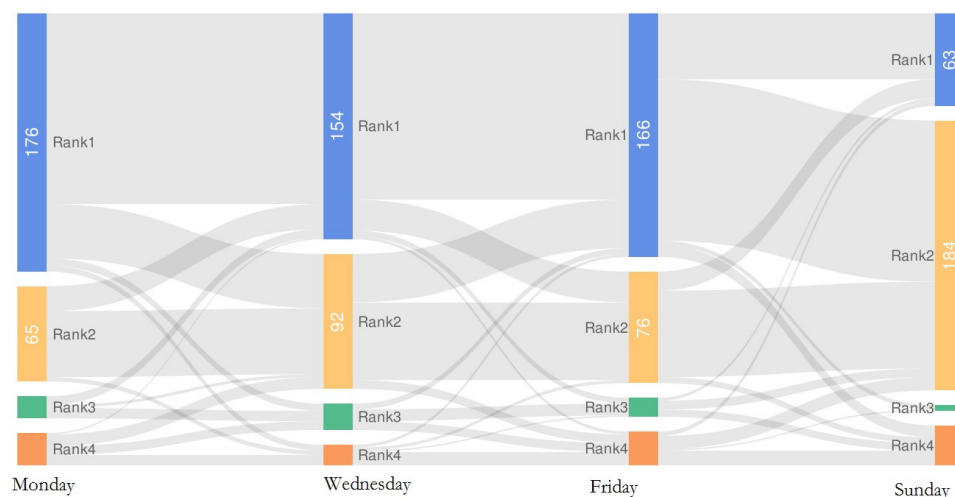


Figure 8. The predictable level of short-term station boarding passenger flow changes in different weeks.

Spatial Distribution of Stations in Different Ranks

Figure 9 shows the spatial distribution of stations in different ranks and weeks. From the distribution on weekdays (10-min interval), it can be seen that the stations in Rank 1 are relatively evenly distributed. Most of the stations in Rank 2 located in the urban areas within the Fifth Ring Road; there are fewer stations in Rank 3 and Rank 4. According to the distribution on weekends (60-min interval), the majority of the stations are in Rank 2, and most of them are located in the Fourth Ring Road. Due to the fact that the percentages of stations in Rank 1 and Rank 2 on weekdays and weekends exceed 85%, the spatial distribution pattern is not significant. Stations in Rank 1 are distributed relatively even in the entire network, whereas stations in Rank 2 are mainly located in the urban areas.

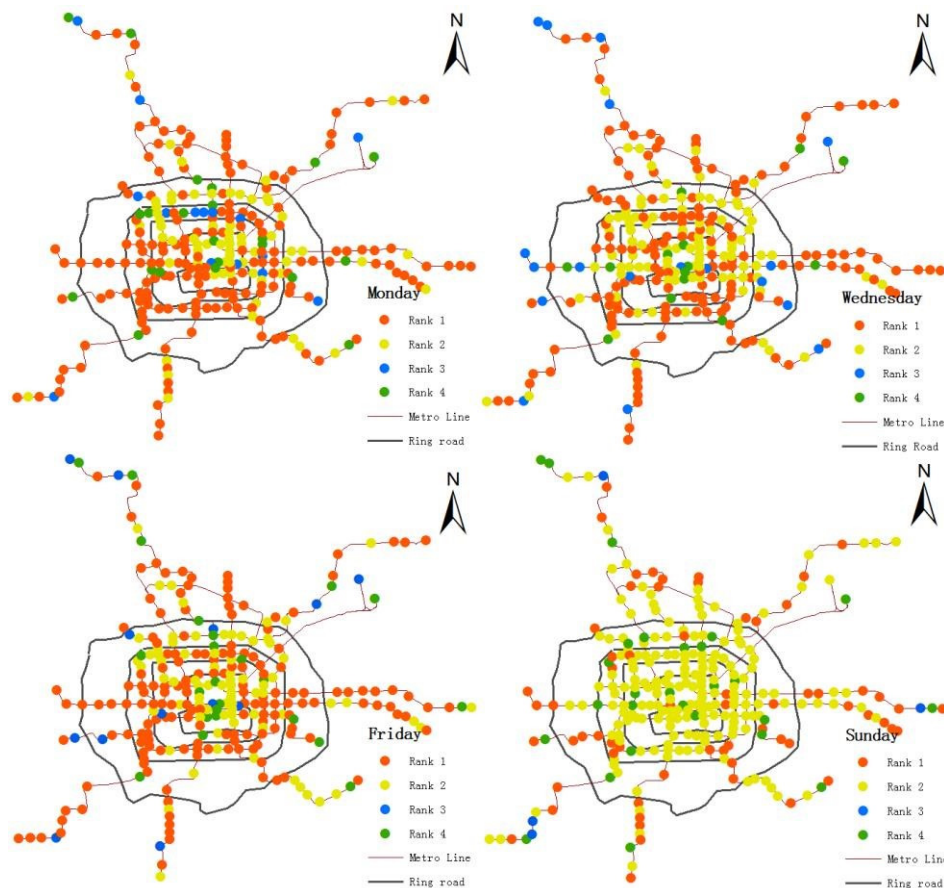


Figure 9. The spatial distribution of stations at different predictable levels.

5.2. Forecasting Passenger Volumes in Different OD Pairs

5.2.1. Comparison of Metrics for Weekdays and Weekends

For each OD, selecting 10-min interval for both weekdays and weekends, we compared the applicability of the two forecasting methods. Since the OD passenger flow has an overall lower similarity, threshold is set at 0.75. In Figure 10, the red-shaded area shows passenger volumes in OD pairs with a higher similarity ($RR_{D_At} > 0.75$), therefore the ODs that fall in this area are suitable for predicting with the same period method. The blue shaded area shows the OD passenger flow with lower variability ($P < 0.05$), thus OD pairs in this area can be predicted using the time series method. The intersection area indicates that both methods are applicable, while the area outside of the shaded region means that using either of the two methods alone cannot achieve accurate prediction results.

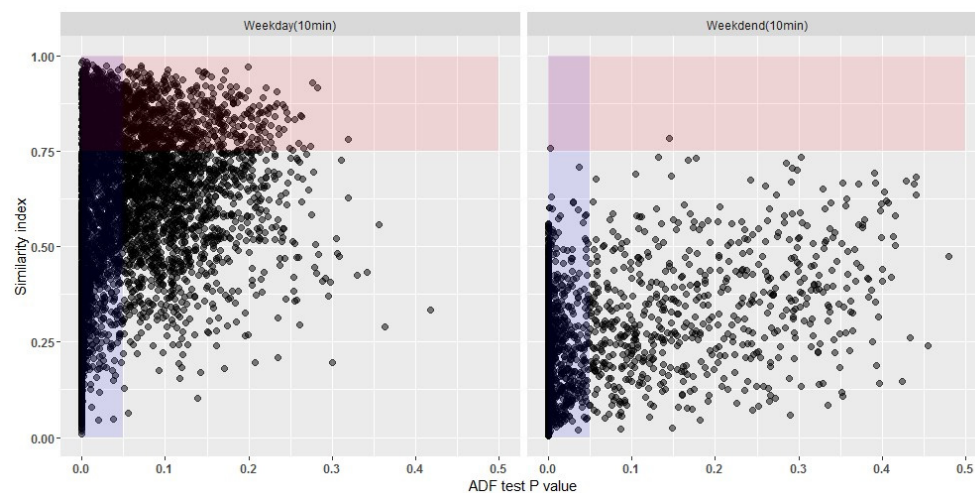


Figure 10. Comparison between the same period and time series predictions of OD flow on weekdays and weekends.

5.2.2. Predictability Level of Passenger Volumes in Different OD Pairs

Ranking the predictivity level of the short-term OD pairs, the result is in Table 4.

Table 4. The level of the predictivity level of the short-term OD pairs.

Level	Feature
Rank 1	ODs that are suitable for both methods. The same period similarity $RR_{D_Δt} > 0.75$ and the time series $P < 0.05$ for the ADF test. These ODs are in the overlapping area of the blue and red shade in Figure 10.
Rank 2	ODs that are more suitable for the same period method. The same period similarity $RR_{D_Δt} > 0.75$ and the time series $P > 0.05$ for the ADF test. These ODs are in the area that is covered by the red shade alone.
Rank 3	ODs that are more suitable for the time series method. The same period similarity $RR_{D_Δt} < 0.75$ and the time series $P < 0.05$ for the ADF test. These ODs are in the area that is covered by the blue shade alone.
Rank 4	ODs that are not suitable for either of the methods. The same period similarity $RR_{D_Δt} < 0.75$ and the time series $P > 0.05$ for the ADF test. These ODs are in the unshaded area in Figure 10.

Changes in OD Ranks on Different Days of the Week

Figure 11 shows variations in the predictable rank of the same OD on different days of a week. Although each rank of OD varies day by day, the predictable ranks of most ODs remain constant during the weekdays, and the number of ODs per rank remains relatively stable. For example, from Monday to Wednesday, although the number of ODs changing from Rank 1 to Rank 2 and from Rank 3 to Rank 4 is relatively high, there are also ODs changing from Rank 2 to Rank 1 and Rank 4 to Rank 3. ODs with constant ranks for three weekdays accounted for 43.1%, of which the ODs belonging to Ranks 1 to 4 accounted for 10.28%, 3.96%, 18.04%, and 10.82%, respectively, indicating that the short-term OD passenger flow is more suitable to the time series method for prediction. On weekends, the majority of all ranks of OD change to Rank 3, and the number of ODs for Rank 3 reaches 4385 pairs, while the number of ODs for Rank 1 and 2 is almost zero. It is worth noting that 85.30% of Rank 4 ODs rises to Rank 3 from weekday to weekend. The OD pairs with less predictability on weekdays, can be predicted on weekends with the time series method, and a few OD pairs transfer from Rank 1 to Rank 4 on weekends. For different days and different ranks of OD, it is important to select the appropriate prediction methods to get better results.

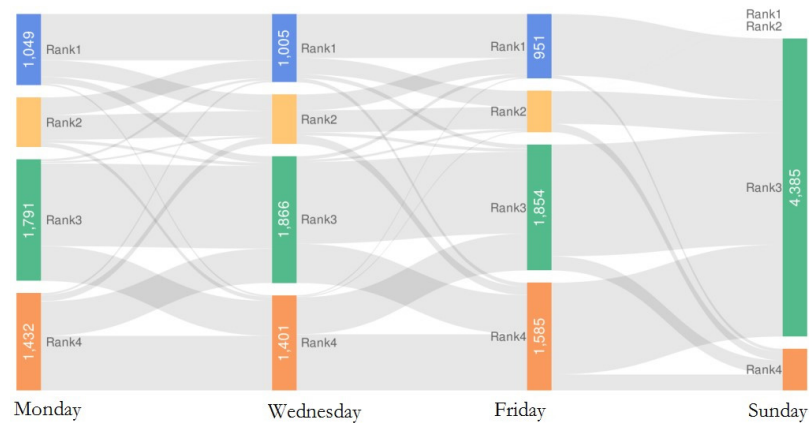


Figure 11. The predictable level of OD passenger flow changes in different weeks.

Spatial Distribution for ODs in Different Ranks

Figure 12 shows the spatial distribution pattern of different ranks of OD in different weeks. The ODs of Rank 1, Rank 2, and Rank 3 are highlighted on three chosen weekdays. On Sunday, OD pairs mainly belong to Rank 3 and Rank 4, thus, Rank 4 OD pairs are highlighted. On weekdays, the OD spatial distribution of Rank 1 shows that there are more ODs outside of the Fifth Ring Road. The ODs in Rank 2 have similar distribution to the ones in Rank 1, but are located in a smaller range. The main reason is that passenger flow is relatively high at these stations. And the number of ODs going from these stations to other stations on the entire network and arriving at these stations is also relatively large. Besides, these ODs have a large space span and a long travel distance. In contrast, the OD in Rank 3 has a smaller spatial range, and locates mainly in the area of the Fourth Ring Road. ODs with long travel distances generally have a station at one end that serves a suburban residential area and customers are daily commuters. Such ODs have higher same period similarity, while short-distance ODs mainly consist of central stations and have a higher stability. On weekends, ODs in Rank 3 and 4 account for the vast majority. ODs in Rank 4 mainly located at the Beijing South Railway Station, Tiantongyuan Station, Huilongguan Station, Xi'erqi Station, Xidan Station, etc. The ODs composed by these stations have large spatial spans and long travel distances, but their similarity and stability are both low on weekends.

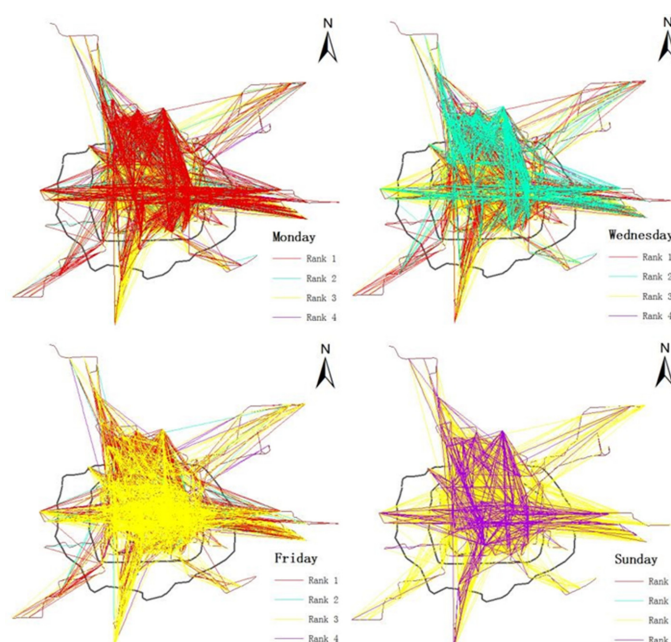


Figure 12. The spatial distribution of OD at different predictable levels.

6. Conclusions

The overcrowding problem is quite common for mass rail transit systems, especially in China with such a large population. To ensure rail passenger safety and mitigate crowding, real-time passenger management measures are often adopted, such as passenger control or train deliver adjustment. To manage passengers in a real-time sense, operators have to know the ridership in advance, such as 15 min earlier. To this end, short-term ridership forecasting of travel demand is very widely addressed. However, existing works focus more on the methodologies of predicting using historical ridership data, and various models are developed for general application. These models may work well for some rail stations but not for others, or achieve good results at one time resolution but bad for others. Thus, the possible forecasting method and time resolution choice with accuracy, to some extent, were investigated.

In this paper we focused on the selection of methods and interval for short-term passenger flow prediction, using the same period method and time series method to analyze the short-term station passenger flow and short-term OD passenger flow, ranking and evaluating the predictability of short-term passenger flow and finally, giving suggestions for stations in different ranks and different time on the selection of interval. The results are listed below:

- (1) The similarity of the short-term station boarding passenger flow increases with the increase of time resolution. On weekdays, the time series stability increases with the increase of time resolution, while on weekends it decreases with the increase of time resolution. The same period similarity of the station boarding passenger flow is much better than the time series stability. It is suggested to choose 10 min and 60 min for weekdays and weekends respectively as the minimum time resolution for forecasting short-term station boarding demand. Stations in Rank 1 and Rank 2 are suitable for the same period method, while stations in Rank 3 are suitable for the time series method, which can make more than 90% of the station's short-term boarding demand of the whole network achieve reliable prediction results.
- (2) The similarity of OD passenger flow increases with the increase of time resolution, while the time series stability decreases with the increase of time resolution. The stability of OD passenger flow is significantly better than similar. It is suggested that the forecast of short-time OD demand is suitable with 10-min interval, and should mainly use the time series method. For weekdays, the same period method should also be used to make predictions. This can provide accurate predictions for about 70% weekdays and 90% weekends OD short-time passenger flow.

Our conclusions reveal that the ridership predicting can be quite case specific and different stations have different travel demand patterns, so does the predicting method. To achieve better results, operators should develop different types of forecasting models for different kinds of stations. Meanwhile, the information hiding in the historical data has its limits, so there is a time resolution cap. One cannot scale down to very high time resolution of forecasting and still assure the required accuracy. This resolution cap can also be case specific.

In the future, researchers and practitioners should develop different categories of forecasting models for short transit ridership predicting according to the ridership characteristics, rather than establish an arbitrary type of model for general application.

Author Contributions: Z.-j.W., modeling guidance and content planning; H.-x.L., manuscript writing and editing; S.Q., modeling guidance and manuscript modification; J.-p.F., data-assisted analysis and mapping guidance; T.W., literature search and review, manuscript writing.

Funding: This research is supported by National Natural Science Foundation of China (No. 51608109) and National Natural Science Foundation of China (No. 51978044). The authors also appreciate the support of Open Foundation of Key Laboratory of Advanced Public Transportation Science, Ministry of Transport, PRC.

Acknowledgments: This research is supported by National Natural Science Foundation of China (No. 51608109) and National Natural Science Foundation of China (No. 51978044). The authors also appreciate the support of Open Foundation of Key Laboratory of Advanced Public Transportation Science, Ministry of Transport, PRC.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this work.

Data Availability: The origin and detailed smart card data [Dump File (.dmp)] used to support the findings of this study were supplied by [Beijing Transportation Information Center] under license and cannot be made freely available, for possible issues of security and personal privacy. Requests for access to these origin data should be made to [Bo WANG, Email: l630680883@126.com].

References

1. Lu, K.; Han, B.; Lu, F.; Wang, Z. Urban rail transit in China: Progress report and analysis (2008–2015). *Urban Rail Transit*. **2016**, *2*, 93–105. [\[CrossRef\]](#)
2. Tsai, T.-H.; Lee, C.-K.; Wei, C.-H. Neural network based temporal feature models for short-term railway passenger demand forecasting. *Expert Syst. Appl.* **2009**, *36*, 3728–3736. [\[CrossRef\]](#)
3. Vlahogianni, E.I.; Golias, J.C.; Karlaftis, M.G. Short-term traffic forecasting: Overview of objectives and methods. *Transp. Rev.* **2004**, *24*, 533–557. [\[CrossRef\]](#)
4. Williams, B.M.; Hoel, L.A. Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results. *J. Transp. Eng.* **2003**, *129*, 664–672. [\[CrossRef\]](#)
5. Dougherty, M. A review of neural networks applied to transport. *Transp. Res. Part C: Emerg. Technol.* **1995**, *3*, 247–260. [\[CrossRef\]](#)
6. Castro-Neto, M.; Jeong, Y.; Jeong, M.K.; Han, L.D. AADT prediction using support vector regression with data-dependent parameters. *Expert Syst. Appl.* **2009**, *36*, 2979–2986. [\[CrossRef\]](#)
7. Rojas, I.; Valenzuela, O.; Rojas, F.; Guillén, A.; Herrera, L.; Pomares, H.; Marquez, L.; Pasadas, M. Soft-computing techniques and ARMA model for time series prediction. *Neurocomputing* **2008**, *71*, 519–537. [\[CrossRef\]](#)
8. Zhang, X.; Mao, B.; Wang, Y.; Feng, J.; Li, M. Wavelet Neural Network-based Short-Term Passenger Flow Forecasting on Urban Rail Transit. *Telkomnika Indones. J. Electr. Eng.* **2013**, *11*, 7379–7385. [\[CrossRef\]](#)
9. Sun, Y.; Leng, B.; Guan, W. A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system. *Neurocomputing* **2015**, *166*, 109–121. [\[CrossRef\]](#)
10. Wei, Y.; Chen, M.-C. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. *Transp. Res. Part C Emerg. Technol.* **2012**, *21*, 148–162. [\[CrossRef\]](#)
11. Jia, Y.; He, P.; Liu, S.; Cao, L. A Combined Forecasting Model for Passenger Flow Based on GM and ARMA. *Int. J. Hybrid Inf. Technol.* **2016**, *9*, 215–226. [\[CrossRef\]](#)
12. Yao, X.; Zhao, P.; Yu, D. Short-time Passenger Flow Origin-destination Estimation Model for Urban Rail Transit Network. *J. Transp. Syst. Eng. Inf. Technol.* **2015**, *2015*, 149–155. (In Chinese)
13. Du, Y.; Snu, Y.; Chen, G. Time resolution Selection for Expressway OD Realtime Prediction. *J. Tongji Univ. (Nat. Sci.)* **2016**, *44*, 1553–1558. (In Chinese)
14. Zhong, C.; Batty, M.; Manley, E.; Wang, J.; Wang, Z.; Chen, F.; Schmitt, G. Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing Using Smart-Card Data. *PLoS ONE* **2016**, *2*, e0149222. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Hauke, J.; Kossowski, T. Comparison of values of Pearson’s and Spearman’s correlation coefficients on the same sets of data. *Quaest. Geogr.* **2011**, *30*, 87–93. [\[CrossRef\]](#)
16. Anvari, S.; Tuna, S.; Canci, M.; Turkay, M. Automated Box–Jenkins forecasting tool with an application for passenger demand in urban rail systems. *J. Adv. Transp.* **2016**, *50*, 25–49. [\[CrossRef\]](#)
17. Oh, C.O.; Morzuch, B.J. Evaluating time-series models to forecast the demand for tourism in Singapore: Comparing within-sample and postsample results. *J. Travel Res.* **2005**, *43*, 404–413. [\[CrossRef\]](#)



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).