

Research Article

Testing the Generality of a Passenger Disregarded Train Dwell Time Estimation Model at Short Stops: Both Comparison and Theoretical Approaches

Dewei Li ¹, Yonghao Yin ¹, and Hong He ²

¹State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

²Guangzhou Railway Polytechnic, Guangzhou, Guangdong 510430, China

Correspondence should be addressed to Dewei Li; dwli.bjtu@163.com

Received 15 June 2017; Revised 1 November 2017; Accepted 11 January 2018; Published 15 February 2018

Academic Editor: Zhi-Chun Li

Copyright © 2018 Dewei Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Train dwell time estimation is a critical issue in both scheduling and rescheduling phases. In a previous paper, the authors proposed a novel dwell time estimation model at short stops which did not require the passenger data. This model shows promising results when applied to Dutch railway stations. This paper focuses on testing and improving the generality of the model by two steps: first, the model is tested by applying more independent datasets from another city and comparing the estimation accuracy with the previous Dutch case; second, the model's generality is tested by a theoretical approach through the analysis of individual model parameters, variables, model scenarios, and model structure as well as work conditions. The validation results during peak hours show that the MAPE of the model is 11.4%, which is slightly better than the results for the Dutch railway stations. A more generalized predictor called “dwell time at the associated station” is used to replace the square root term in the original model. The improved model can estimate train dwell time in all the investigated stations during both peak and off-peak periods. We conclude that the proposed train dwell time estimation model is generic in the given condition.

1. Introduction

Train dwell time estimation is a critical issue for both the scheduling and rescheduling phases. In the scheduling phase, accurate estimation of the train dwell time provides necessary inputs to the timetable and allows the timetable to match the passenger demand. In the rescheduling phase, the estimation results of the train dwell time can be used to predict potential conflicts between train lines.

To estimate dwell time accurately, the development of a dwell time estimation model is necessary. One of the most critical issues for broad applications of the train dwell time estimation model is its generality. The generality of a model can be defined as a model that is generic rather than a more specific or detailed one [1]. In other words, a generalized dwell time estimation model can be used to estimate dwell times in different instances with acceptable accuracy.

To test the generality of a train dwell time estimation model, there are two possible approaches. One intuitive

methodology is to apply the model to more or a wider range of independent cases; if the accuracies under these different cases are acceptable, then the model can be classified as a relatively general model. This approach is called a comparison approach. However, comparison approaches may face problems like how many cases they should cover and what the wider range means. Another methodology is to analyze the parameters and variables of the model theoretically to prove its generality [1], which is called a measure theoretical approach. The generality of a model is the measure assigned to the set of possibilities to which it applies. Such a method can provide an account of the minimum formal machinery required to make precise, quantitative assessments of generality.

The main methodology of this paper is as follows: first, a train dwell time estimation model is selected; second, the generality of the selected model is analyzed by using both comparison and theoretical approaches. In the comparison

TABLE 1: Main features of existing dwell time estimation models.

Source	Passenger traffic dependency	Model methodology	Validation		Railway mode	Station type
			Stations and lines	Country		
Lam, 1988	Pr	Lr	3S on 3L	HKG	Subway	Ich
Lin, 1992	Pr	NLr	2S on 1L	USA	Light rail	—
Puong, 2000	Pr	NLr	2S on 1L	USA	Subway	Ich and Ime
Buchmueller et al., 2008 [16]	Pr	Di	—	SWE	Hsr	—
Zhang et al., 2008 [3]	Pr	Mi	3S on 3L	CHN	Subway	Ich and Ime
Hansen, 2010	Pd	Lr	2S on 1L	NED	Hsr	Ime
Jong, 2011	Pr	Lr	1S on 1L	TPE	Hsr	Ich
Sourd, 2011	Pr	Mi	1S on 1L	FRA	Hsr	—
Yamamura, 2013	Pr	Mi	—	JPN	Subway	—
Kecman, 2015	Pd	Lr	—	NED	Hsr	Ich and Ime
Seriani, 2015	Pr	Mi	—	—	Subway	—
Hänseler et al., 2015 [17]	Pr	Mi	1S on 1L	SWE	Hsr	Ich
Li, 2016	Pd	NLr	1S on 1L	NED	Cr	Ime
This paper	Pd	NLr	1S on 1L and 7S on 1L	NED and CHN	Cr and subway	Ich and Ime

Note. Pr: passenger regarded; Pd: passenger disregarded; Lr: linear regression; NLr: nonlinear regression; Mi: microscopic simulation; Di: distribution; 3S on 3L: 3 stations on 3 lines (other abbreviations et cetera); HKG/USA/CHN and others: the acronym of countries; Hsr: high-speed railway; Cr: commuter railway; Ich: interchange station; Ime: intermediate station.

approach, an independent dataset is applied, and the estimation result is compared with the previous one. In the theoretical approach, the model is analyzed from four aspects: variable analysis, parameter analysis, scenario analysis, and structure and condition analysis. Finally, based on the testing of generality, an improvement is made to the model.

The remainder of this paper is organized as follows. In Section 2, a literature review is presented. Section 3 introduces the selection of the train dwell time estimation model. In Section 4, the comparison approach is applied to the selected model. A new dataset from the Beijing urban rail transit is applied and then compared to the original result, which had been validated by data from Dutch railway stations [2], and the estimation accuracy is analyzed by comparing the results of the models. In Section 5, the train dwell time estimation model is analyzed by the theoretical approach. Section 6 improves the generality of the proposed model. In Section 7, three controversial issues are discussed to explain the condition of the model. Section 8 presents several conclusions.

2. Literature Review

In the literature, many types of train dwell time estimation models have been proposed. The existing models can be classified according to different criteria such as passenger traffic dependency, modeling methodology, validation, railway mode, and type of station, all of which are listed in Table 1.

From Table 1, there are two distinct models that are interrelated and can be distinguished from the research

direction. One is a passenger regarded model and the other is a passenger disregarded model.

In the passenger regarded model, the dwell time is estimated based on the alighting and boarding time of the passengers. The numbers of boarding and alighting passengers are the key inputs of this kind of model. The approaches of the passenger regarded train dwell time estimation mainly include two types: a microscopic simulation model and a regression model.

Among the microscopic simulation models, these studies focused on the microlevel movement during alighting and boarding processes by simulation modeling. Most of the microscopic simulation studies are based on field data collected in test stations and lines. They are used to explain the effect of the uncertainties of passenger behaviors on dwell times. Zhang et al. [3] presented a cellular automata-based alighting and boarding microsimulation model for passengers to estimate the dwell time, which considers individual desired speed, pressure from passengers behind, personal activity, and behavioral tendencies (dependent on gender, age, etc.). Seriani and Fernandez [4] determined the effect of passenger traffic management in the boarding and alighting time at metro stations by simulations and experiments. Yamamura et al. [5] developed a simulator to estimate the dwell times of trains when various improvements for reducing dwell time were taken. The simulator was based on a multiagent model and simulated each passenger's behavior both in a train and on a platform. Sourd et al. [6] presented a behavioral based simulation model dedicated to

the evaluation of the dwell time for trains, which separate the environmental model and can be computed easier and faster. The results of these studies showed that passenger behaviors vary from different countries or districts because of diversity characteristics and service levels and other factors that may influence pedestrian movement. These models consumed a lot of time during the simulation phase and their inputs were varied; thus they can be hardly used in real-time applications.

Among the passenger regarded regression models, the train dwell times have usually been estimated using a linear or nonlinear function of the number of alighting and boarding passengers. Lam et al. [7] and Wirasinghe and Szplett [8] both presented linear regression models to estimate the dwell time, with the difference being that the former assume that the boarding passengers on the platform take on a uniform distribution and the latter consider a nonuniform distribution of the boarding passengers. Lin and Wilson [9], Parkinson and Fisher [10], and Puong [11] developed a nonlinear estimation model, and they considered the number of standing passengers in the vehicle and their interactions with boarding and alighting passengers. These models were usually validated for specific rolling stocks [12, 13], stations [14], and the number of alighting and boarding passengers [15]. Specifically, Harris [12] found that the interior layout of the train typically improves the dwell time estimation model accuracy. Jong and Chang [13] estimated the alighting and boarding times at a specific station as a function of the numbers of alighting and boarding passengers and different train services that implied the influence of rolling stock. Xenia and Nick [14] discussed the influence of step height between the train and platform and the door width to the dwell time estimation. Daamen et al. [15] found that the number of boarding and alighting passengers was the main determinant of the dwell times. The problems with these models were that many background variables were not included, such as the composition of the passenger population configuration of the rolling stock and the type of station (for details, refer to [2]), which had an irrefutable impact on the dwell time. However, none of the existing models fully took all of the influencing factors into account.

Among the passenger disregarded models, the dwell time was estimated using many variables that were not related to passenger demand. To the best of our knowledge, so far, regression is the only approach that has been used in passenger disregarded models. Hansen et al. [18] and Kecman and Goverde [19] estimated the train dwell time as a function of its arrival delay, which was derived from the track occupancy data of the Dutch railway. Li et al. [2] proposed a train dwell time estimation model that improved the accuracy of these models. This model predicted the train dwell time using several substitute variables, including the actual dwell time of consecutive trains at the target station and the dwell time of the target train at consecutive stations, which could indirectly reflect the passenger demand [2]. However, the generality of the models was not validated for cases from other countries. Meanwhile, the result of the research was not good enough for an estimate model. There is still space to improve the accuracy of the dwell time estimation regression model.

With regard to the validation of the existing models, most of the former studies included only a few stations belonging to one railway (as shown in Table 1). Instead of validating the model in one specific railway in most of the former works, this paper focuses on two different independent cases: a commuter railway case in Netherlands and a subway case in China. However, despite the lack of a theoretical approach for all the existing models, this paper bridges the gap by providing a general framework for testing the generality of a dwell time model through theoretical analysis.

3. Model Description

To select a possible generic train dwell time estimation model, both the input data availability and model characteristics should be considered. We noticed that all the passenger regarded models could not be selected because the data, such as the number of boarding and alighting passengers, were usually not available, especially in real time. However, the generality of the passenger disregarded model, especially the model proposed by Li et al., is worthy of discussion because the variables considered in Li et al.'s model do not depend on specific types of rolling stocks and traffic conditions [2]. Therefore, this paper analyzes the extent to which the model can be generalized by both comparison and theoretical approaches.

Li et al.'s model studied the dwell time estimation problem particularly for short stops without passenger demand by means of a statistical analysis of track occupation data from Netherlands. In their paper, the factors of influence on dwell time are classified into five categories: passenger, rolling stock, station, operation, and external factors. The five categories include the majority of influencing factors. Then, with the analysis of these factors, 10 potential predictors, including time variation, length of target train, length of preceding train, departure delay at previous station, dwell time of target train at previous station, second previous station, dwell time of preceding train at target and previous station, and dwell time of the same train during the last week, are selected. Various combinations (including linear items and nonlinear items) are tried to establish 10 different parametric regression models. With the correlation analysis and experimental testing, the model shown in formula (1) was finally developed; this model was found to perform the best among their 10 regression models. For details, refer to Li et al., 2016.

$$DT_k^s = c + \beta_1 L_k^s + \beta_2 L_{k-1}^s + \beta_3 DT_{k-1}^s + \beta_4 \sqrt{DT_k^{s-1} * DT_k^{s-2}}, \quad (1)$$

where DT_k^s indicates the dwell time of train k at station s , L_k^s indicates the length of train k at station s , L_{k-1}^{s-1} indicates the length of train k at station $s - 1$, DT_{k-1}^s indicates the dwell time of the previous train at the target station, and DT_k^{s-1} and DT_k^{s-2} indicate the dwell time of the target train at the previous station and at the second previous station, respectively. c , β_1 , β_2 , β_3 , and β_4 are the parameters of the variables.

The reason why this model can work well in practice lies in the selected predictors. Unlike other models, Li et al.'s model does not consider the station layout, rolling stock configuration, or passenger behavior directly because these factors are involved in independent variables in the model and act on the dependent variable indirectly. For example, the predictors of the model are mainly the dwell time of the preceding train DT_{k-1}^s and the dwell time at the previous station DT_k^{s-1} . The common characteristic of the preceding train and the target train is that both trains stopped at the same station, such that parameter β_2 can reflect the effect of the station. The common characteristic of the dwell time at the previous station and the dwell time at the target station is that both times are for the same train, such that β_3 can reflect the effect of the rolling stock. Passenger demand appears in two ways: one is simply separating the peak and off-peak hour models, and the other is to reflect the demand difference via the different lengths of the trains. It should be noted that the variance of passenger demand between different stations is inevitable and is mainly indicated by parameter c . Other remaining influencing factors are also implied in this parameter. Accordingly, the model avoids the direct input of all the influencing factors, which would otherwise be difficult to obtain. However, to what extent the relationship among these variables holds is not clear. Therefore, the generality of such relationships is worthy of discussion.

To test the generality of the model, in the following two sections, two different approaches are used. First, a Beijing dataset is applied to the model that was established based on the Dutch railway station, and the estimation accuracy is analyzed by comparing the two results. Because the two datasets are independent in both time and space, it can be inferred that the validated result reflects the generality of the model to some extent. Second, the quantitative generality of the model is analyzed by a theoretical approach.

4. Generality Analysis by Comparison Approach

The original dataset for Li et al.'s model development is from Dutch railway stations. To test whether it is still applicable for a totally independent case, a new dataset, which is dependent on the original validated case, is selected from the Beijing urban railway. First, an empirical study is conducted by analyzing the difference between the scheduled and actual dwell time in the Beijing urban rail transit, while the distributions of the actual dwell time are also analyzed. Then, these results are compared with the results from Netherlands. Second, the dataset of the Beijing urban rail stations is used to regress the original model that was previously validated for a station in the Dutch railway system, and the regression parameters and the accuracy of the results are compared with the proposed model. In this way, the generality of the proposed model is analyzed.

4.1. Dataset Resources

4.1.1. Dataset Collection. To study the generality of the model by a comparison approach, datasets from China are carefully

TABLE 2: The selected stations and their classifications.

Number	Target station	Classification
1	Renmin University	Intermediate station
2	Weigongcun	Intermediate station
3	National Library	Interchange station
4	Beijing Zoo	Intermediate station
5	Xizhimen	Interchange station
6	Xinjiekou	Intermediate station
7	Ping'anli	Interchange station

selected. To obtain precise datasets from the Beijing railway stations, a line that has been operating for a relatively long amount of time is selected so that the passenger demand is stable. Thus, line four of the Beijing urban rail transit is selected for this empirical study; the stations of this line are shown in Figure 1.

This line opened in 2009, has a length of 50 km, and comprises 35 stations; all trains stop at every station, and no train overtaking occurs at any station. In this railway, Anheqiao North is the first station, and Tiangongyuan is the last station, where the trains are able to turn around. The interchange stations are Haidianhuangzhuang, National Library, Xizhimen, Ping'anli, Xidan, Xuanwumen, Caishikou, Beijing South Railway Station, and Jiaomen West, where passengers can transfer to another line; however, passenger connections are not considered in the timetable, so we neglect the "adhere to schedule effect" which may exist in some transfer stations in other cases. The remaining stations are intermediate stations.

In this paper, four intermediate stations and three interchange stations are selected, and all of the selected stations are consecutive. The selected stations are listed in Table 2.

There are two different ways to obtain the dwell time for the two cases. In the Dutch case, almost three months of data were used to analyze the distribution of the actual dwell time in different periods. The dwell times at the selected stops and trains were estimated based on the track occupation data. In Netherlands, track occupation data were collected using a train describer system (TROT), which provided the exact time of occupation and clearance of track sections [20]. By using a dwell time estimation algorithm [21], a total of 17,306 trains running from 1 September 2012 to 30 November 2012 were processed and analyzed. The sample was separated from the off-peak period ($[9:00, 16:00] \cup [18:30, 20:00]$) and the peak period ($[6:30, 9:00] \cup [16:00, 18:30]$). The peak period was further classified as a morning peak period ($[6:30, 9:00]$) and an afternoon peak period ($[16:00, 18:30]$). However, it is difficult to automatically obtain the actual arrival and departure time of the trains in the Beijing urban rail transit system, such that it was necessary to investigate manually the dwell times of selected stations. The dwell times were investigated for stations and trains. At first, the investigators were distributed among consecutive stations from Anheqiao to Xuanwumen station, on 25 September 2015, which is a workday, from 7:28 am to 11:11 am to include data from both the peak hours $[7:28, 9:00]$ and off-peak hours $[9:00, 11:11]$. In

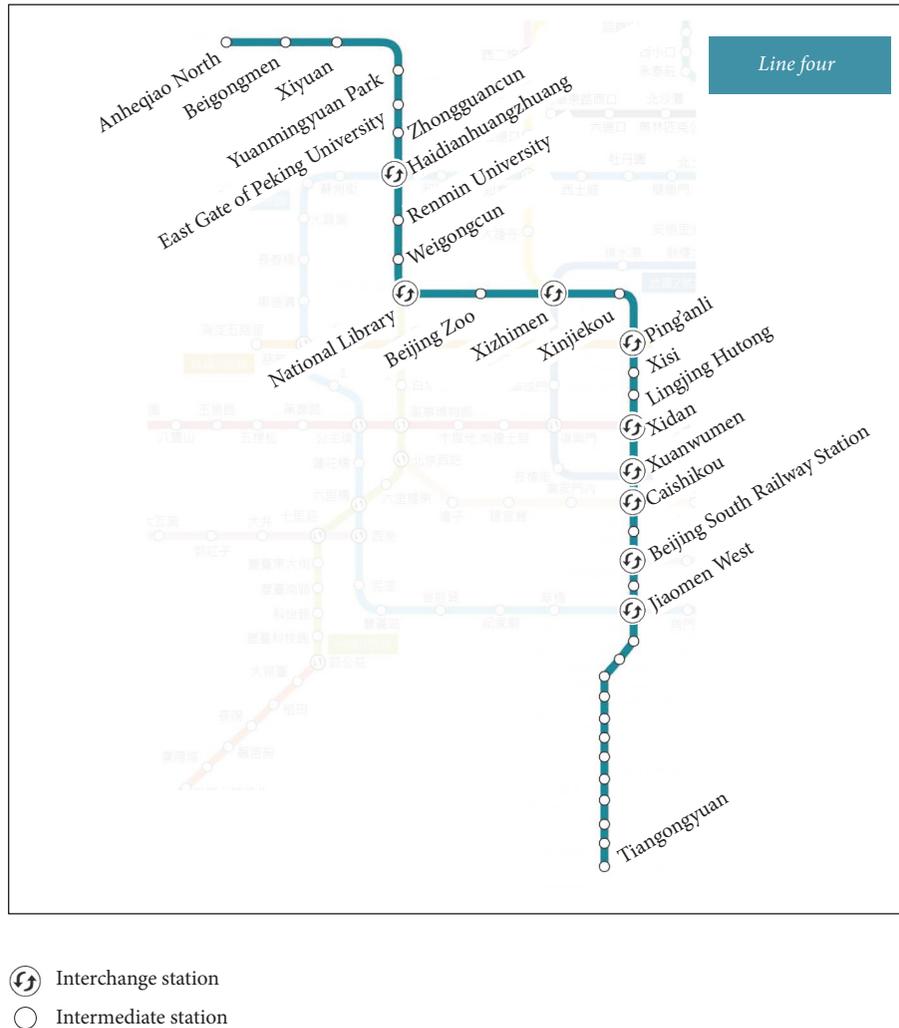


FIGURE 1: Stations of line four of the Beijing urban rail transit system.

total, 71 train arrival and departure times were recorded by using a stop watch. Second, the investigators were requested to board the same number of trains successively departing from the start station of the line, and each investigator recorded the dwell time of the train at every station along the line in one direction by a stop watch. The dwell time is defined as the time from the moment the train stops at the station to the moment that the train begins to leave. The scheduled timetable for line four was also obtained from the operating company. Using the original recorded dataset, we are now able to calculate the dwell time of every train at every station.

4.1.2. Dataset Comparison between the Actual and Scheduled Dwell Times. The original recorded datasets of the investigated stations in Beijing are processed and analyzed, and we find that the actual and scheduled dwell times are significantly different, whether by more or less time. The analysis results are shown in Table 3. The maximum dwell time is longer than the scheduled dwell time by more than 20 s at every station. The minimum dwell time is shorter than the scheduled dwell time by no more than 10 s at most of stations. In addition, the

average dwell time is longer than the scheduled dwell time at every station. We find that the actual dwell time fluctuates significantly, and the difference between the maximum and scheduled dwell times is larger than the difference between the minimum and scheduled dwell times. Next, we compare the datasets between the two cases.

4.1.3. Dataset Comparison between the Two Cases. In this section, the datasets from two countries are compared. At first, the two datasets are statistically described; and then the dwell time distributions and the differences between the actual and scheduled dwell times of the two cases are analyzed.

(1) Data Statistical Descriptions. The average, minimum value, maximum value, standard deviation, and variance of the two datasets are compared. Moreover, the datasets in the two cases are compared by applying the two-sample Kolmogorov-Smirnov test (K-S test) to test whether the two datasets are significantly different. The result is listed in Table 4. This table shows that the two datasets are significantly different. Thus, the two datasets can test the generality of the model.

TABLE 3: Comparison of the actual and scheduled dwell times of the investigated stations.

Stations	Max dwell time	Min dwell time	Average actual dwell time	Scheduled dwell time	Average absolute error	Average relative error	Root mean square error	Max absolute error
Renmin University	69 s	25 s	34.76 s	30/35 s	5.51 s	17.69%	9.31 s	39 s
Weigongcun	67 s	25 s	35.76 s	30/35 s	6.2 s	19.94%	10.42 s	37 s
National Library	69 s	29 s	46.55 s	35 s	12.59 s	35.98%	15.5 s	34 s
Beijing Zoo	65 s	24 s	43.37 s	40 s	8.55 s	21.37%	10.65 s	25 s
Xizhimen	67 s	33 s	46.62 s	45 s	6.38 s	14.18%	8.36 s	22 s
Xinjiekou	70 s	27 s	38.68 s	35 s	6.1 s	17.42%	11.62 s	35 s
Ping'anli	95 s	24 s	45.45 s	35 s	11.77 s	33.64%	16.72 s	60 s

TABLE 4: Comparison between Dutch and Beijing datasets.

Dataset	Average	Minimum	Maximum	SD	Variance	K-S test
D_peak	52.67 s	40 s	71 s	10.28 s	105.75	0.375
B_peak	47.68 s	29 s	65 s	8.99 s	80.84	
D_off-peak	41.67 s	17 s	77 s	16.46 s	271.06	0.505
B_off-peak	38.68 s	24 s	65 s	9.38 s	88.04	

Note. "D" denotes Dutch case; "B" denotes Beijing case; "SD" denotes standard deviation; "K-S test" denotes Kolmogorov-Smirnov test.

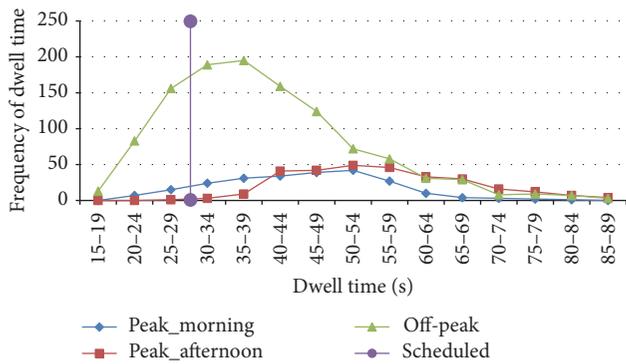


FIGURE 2: Distribution of the actual and scheduled dwell times in the Dutch railway stations.

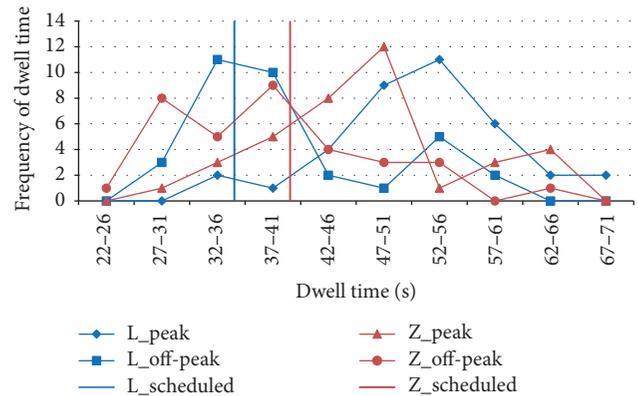


FIGURE 3: Distribution of the actual and scheduled dwell times in the Beijing urban railway stations.

(2) *Dwell Time Distributions.* Because the distribution of the actual dwell time varies significantly during different periods [21], previewing the data and studying the generality of the model through a comparison require datasets from the two countries to be compared. The dwell time distributions of the two cases during peak and off-peak hours are compared, as shown in Figures 2 and 3, respectively.

In the Dutch case, as shown in Figure 2, the maximum frequency value of the actual dwell time (broken line) during both the peak and off-peak hours was higher than the scheduled dwell time (vertical line) because the actual passenger demand could not be satisfied. Therefore, to match the actual passenger demand, the scheduled dwell times were extended.

In the Beijing case, the dwell time is collected manually by investigators. The distribution of the dwell time is analyzed for two typical types of stations during both peak and off-peak hours. The National Library interchange station and the

Beijing Zoo intermediate station are selected. The result is shown in Figure 3. The vertical line indicates the scheduled dwell time, and the broken line denotes the actual dwell time frequency distribution. The maximum frequency value of the actual dwell time during the peak period is higher than that during the off-peak period at both the Beijing Zoo station (Z_off-peak) and National Library station (L_off-peak). The dwell time at the National Library (L_peak) interchange station is larger than that of the Beijing Zoo (Z_peak) intermediate station during peak hours. Conversely, during off-peak hours, the dwell time at the interchange station is less than that of the intermediate station. The primary reason for the greater dwell time is that the passenger demand at the Beijing Zoo station is higher than National Library station; however, the number of transfer passengers is high during the peak period, which also explains why the scheduled dwell time for the Beijing Zoo station is larger

than that of the National Library station. Furthermore, the maximum frequency value of the actual dwell time during the peak hours is higher than the maximum value during off-peak hours in both the Beijing and Dutch railway stations. That is, compared with the dwell time in the off-peak hours, the dwell time is likely to increase during peak hours.

In summary, the distributions of the actual dwell time in the Dutch railway stations and in the Beijing urban rail transit have some common characteristics: regarding the maximum frequency value, the actual dwell time during the peak period is longer than that during the off-peak period. Meanwhile, the maximum frequency values in the two cases are similar. For example, during the peak period, the maximum value of the actual dwell time is 47–51 s at the Beijing Zoo station and 52–56 s at the National Library station, which is similar to the dwell time of 50–54 s during the morning peak period in the Dutch railway stations.

(3) *Difference between the Actual and Scheduled Dwell Times.* In general, the train drivers considered in this study strive to drive in accordance with the timetable. However, differences between the scheduled and actual dwell times do occur, making it necessary to analyze the difference between the scheduled and actual dwell times in both the Dutch railway stations and the Beijing railway stations.

The time period is separated into the peak period and the off-peak period: the peak period lasts from 7:30 to 9:00 am, and the off-peak period lasts from 9:00 to 10:30 am. Two typical stations, the National Library interchange station and the Beijing Zoo intermediate station, are selected to analyze the differences between the scheduled and actual dwell times during the peak and off-peak hours, respectively.

Figure 4 compares the actual and scheduled dwell times in the Dutch railway stations during different periods of a workday on Dutch railway stations. The different line types in the figure represent the dwell time during different periods. Most of the actual dwell time is greater than the scheduled dwell time because the passenger demand is so high that the scheduled time cannot meet the actual demand.

Figure 5 compares the actual and scheduled dwell times of the two stations during peak and off-peak hours in the Beijing railway stations. The different line types in the figure represent the different stations and the dwell time during different periods. During the peak hours, most of the actual dwell times are clearly larger than the scheduled dwell times. In contrast, during the off-peak hours, the actual and scheduled dwell times are closer, and the actual dwell time may even be shorter than the scheduled dwell time because the passenger demand during the off-peak hours is less than that during the peak hours. Furthermore, the dwell time of the National Library interchange station is longer than that of the Beijing Zoo intermediate station during peak hours; conversely, during off-peak hours, the dwell time of the interchange station is less than that of the intermediate station. The primary reason for this result is that the number of transfer passengers is high during the peak period; however, the passenger demand at the Beijing Zoo station is higher than that of the National Library station during the off-peak period, which also explains why the

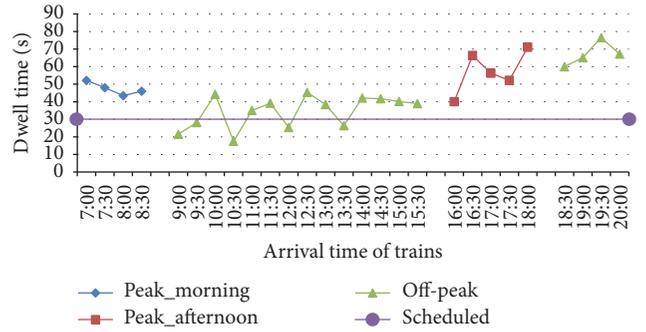


FIGURE 4: Comparison of the actual and scheduled dwell times in the Dutch railway stations.

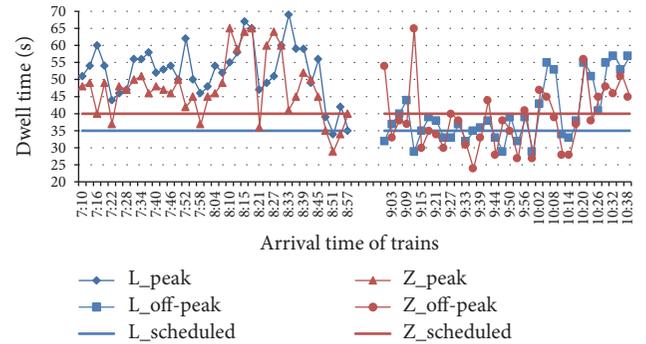


FIGURE 5: Comparison of the actual and scheduled dwell times in the Beijing railway stations.

scheduled dwell time for the Beijing Zoo station is greater than that of the National Library station.

In summary, the differences between the actual and scheduled dwell times in the Beijing urban rail transit system and in the Dutch railway stations have common characteristics: most of the actual dwell times are longer than the scheduled dwell times during the peak period. The actual and the scheduled dwell times are closer during the off-peak period, and the actual dwell time is likely to increase during the peak period.

4.2. Generalization Analysis Approach Based on Comparison.

In this section, the generality of the selected train dwell time estimation model is analyzed by the comparison approach. The collected data are separated into two parts, the first of which is used to calibrate the model. The other part is used to validate and measure the error of the model. Ten records are selected randomly from the data as the validation set, and the remainder are used as the learning sample in the calibration part. In other words, the model is validated by applying the model to the Beijing urban railway dataset; the results are compared with the original values. Accordingly, this method verifies whether the model can cover wider scenarios.

4.2.1. *Model Preparation.* First, we consider the predictors that are used in the original model [2]; some small adjustments are made to adapt these parameters to the Beijing dataset. The lengths of the trains on line four of the Beijing

TABLE 5: The correlations between the input variables and the dwell time.

Stations	Periods					
	Peak			Off-peak		
	DT_{k-1}^s	DT_k^{s-1}	DT_k^{s-2}	DT_{k-1}^s	DT_k^{s-1}	DT_k^{s-2}
Renmin University	—	—	—	—	—	0.36
Weigongcun	0.41	0.84	—	—	0.53	—
National Library	0.54	0.54	0.35	0.54	0.36	—
Beijing Zoo	0.46	0.55	—	0.46	0.77	0.56
Xizhimen	0.47	0.49	0.34	—	0.51	—
Xinjiekou	0.73	0.68	0.60	—	0.43	0.48
Ping'anli	0.35	0.59	0.44	—	0.36	—

TABLE 6: Performance measures.

Indicators	Peak period	Off-peak period
Adj- R^2	0.594	0.162
RMSE	7.21 s	6.21 s
MAPE	11.4%	13.1%

Note. “Adj- R^2 ” denotes the adjusted coefficient of determination; “RMSE” denotes the root mean square error; “MAPE” denotes the mean absolute percentage error.

urban rail transit system are the same during the operating period, and the types of all of the rolling stock are the same. However, the Dutch railway stations have different train lengths at the same railway stations. To accommodate this fact, the trains’ length predictors are neglected and transformed into a constant in the regression model. Thus, in this paper, the predictors are DT_{k-1}^s , DT_k^{s-1} , and DT_k^{s-2} .

Second, a simple correlation analysis is performed between the selected predictors and the dwell times at different stations. The result in Table 5 shows that most of the predictors are significantly different from zero ($\alpha = 0.01$) at different stations during different periods. A relatively strong linear relationship exists for Weigongcun during the peak hours and Beijing Zoo during the off-peak hours. Most of the predictors have correlation coefficients larger than 0.5, especially during the peak periods. In the Dutch cases, correlations between these three variables for the merged data of all stations are 0.376, 0.456, and 0.381, respectively. Therefore, the relationship between the variables is weak but still holds even if the instances are different.

Third, the dataset of the Beijing railway stations is used to calibrate the parameters in formula (1). The trains’ length is neglected and treated as a constant. Therefore, to maintain the stability of formula (1), we set $\beta_1 = 0$ and $\beta_2 = 1.6$ for the six-car train based on the Dutch dataset (because the trains’ length on line four in the Beijing urban rail transit is six cars). The parametric regression results for the peak and off-peak periods are presented in formulas (2) and (3), respectively. The models are regressed using the calibration dataset, which includes the predictors for the 7 selected stations in the Beijing railway stations. Formula (2) and formula (3) can predict the dwell time of all 7 stations during the peak and off-peak hours, respectively.

$$DT_k^s = -0.83 + 1.6 \times L_{k-1}^s + 0.76 \times DT_{k-1}^s + 0.04 \times \sqrt{DT_k^{s-1} * DT_k^{s-2}}, \quad (2)$$

$$DT_k^s = 7.02 + 1.6 \times L_{k-1}^s + 0.30 \times DT_{k-1}^s + 0.24 \times \sqrt{DT_k^{s-1} * DT_k^{s-2}}. \quad (3)$$

Three indicators are introduced to evaluate the estimation accuracy of the results: the adjusted coefficient of determination (adj- R^2), the root mean square error (RMSE), and the mean absolute percentage error (MAPE). Adj- R^2 reflects the level to which the regression model is interpreted by the variables: the larger this indicator is, the better the datasets are explained. The other two indicators are taken from the literature [2]; the RMSE reflects the actual error, and the MAPE is used to evaluate the estimation accuracy. The performance measures of the results based on the Beijing dataset are shown in Table 6.

Table 6 shows that the model accuracy for the Beijing case is better and that the accuracy of the adapted model based on the data for the investigated stations of the urban railway line 4 in Beijing is satisfactory, in particular for the peak period.

4.2.2. Parameter Comparison. The parameters of the model that which based on the two datasets are shown in Table 7, which indicates that the parameters in both cases have similar tendencies, such as the constants during the peak period being negative numbers. The different parameter values indicate the contributions of the predictors to the dwell time, as stated in the literature [2]: β_3 reflects the effects of the stations, and β_4 reflects the effects of the rolling stock. When neglecting train length, in the Dutch dataset, the value of β_4 is greater than that of β_3 during the peak period; in other words, the contribution of dwell time at the previous station is positive, whereas the dwell time of previous trains has almost no effect. The variables of the same trains at different stations have a strong relationship, indicating that the rolling stock factor plays an important role in the dwell time of the Dutch railway which is reflected in reality. In our investigation,

TABLE 7: Model parameters in the different datasets.

Parameters	Dutch dataset		Beijing dataset		
	Peak period	Peak period	SD	Off-peak period	SD
c	-13.07 (0.01)	-0.83 (0.00)	0.23	7.02 (0.00)	0.37
β_1	—	—	—	—	—
β_2	1.60 (0.00)	1.60 (—)	—	1.60 (—)	—
β_3	0.03 (—)	0.76 (0.00)	0.04	0.30 (0.00)	0.06
β_4	1.11 (0.00)	0.04 (0.00)	0.05	0.24 (0.00)	0.07

Note. p values are shown in brackets; “SD” denotes standard deviation.

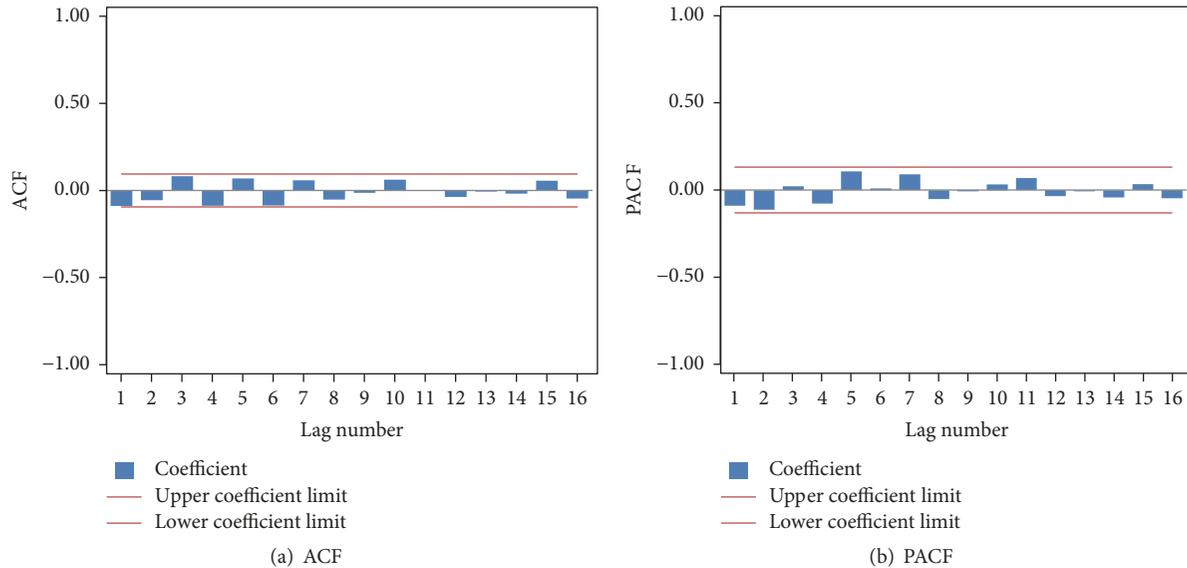


FIGURE 6: Residuals analysis by ACF and PACF.

the yard layout and the sizes of the selected stations in the Dutch railway stations are similar, but the trains' lengths differ significantly; thus, the trains' lengths play a key role in the Dutch model. Conversely, the situation in the Beijing urban rail transit during peak hours differs from that in the Dutch railway stations. The trains' lengths on line four of the Beijing urban rail transit are all the same, but the platform layouts, for example, the positions of the stairs and escalators, and the sizes of the selected stations vary; thus, the station factors play key roles in the model based on the Beijing dataset.

The p values and standard deviation of the predictors are listed in Table 7. It is shown that the p values for all parameters are smaller than 0.05, which indicates that the values of the parameters are significantly different from zero. Meanwhile, we also analyze the residuals based on model diagnostics for the time series. The autocorrelation function (ACF) and partial autocorrelation function (PACF) are shown in Figure 6. The ACF and PACF of the residuals show that the residuals do not correlate statistically and accord with a white noise sequence. In conclusion, the regression model can be used in the Beijing cases.

In addition, the model fails during the off-peak hours for the Dutch case but is still applicable during the off-peak hours for the Beijing case. One possible reason for this is that the scheduled headway is large during the off-peak

hours in the Dutch case, which tends to decorrelate the dwell times. Another reason is that the passenger demand in the Dutch case during the off-peak period is so small that the random factors play a key role in the dwell time. However, the headway (Figure 9) and the passenger demand of the Beijing case during the off-peak period are similar to those of the Dutch case during the peak period, meaning that the influencing factors of the dwell time may be similar. The results show that the values of β_3 and β_4 in the Beijing dataset are small during the off-peak period. Because the variation in the passenger demand is tolerable, the influencing factors have a weak effect on the dwell time; thus, the dwell time at the target station tends to be stable. In summary, the parameters can be justified by analyzing the actual situation.

4.2.3. Performance Comparison. The accuracy of the parametric model based on the Beijing dataset is compared with those of existing models reported in literature (see Table 8). The datasets published in various reports differ from this article. The comparison could have some bias, yet it could reflect the effectiveness and potential of the selected model.

The results from the comparison of the author's model in the different cases are discussed in further detail. First, $\text{adj-}R^2$ is high in the model based on the Beijing dataset during peak hours, which shows that the model can explain most

TABLE 8: Comparison of the regression results for different datasets.

Indicators	Puong [11]	Hansen et al. [18]	Kecman (2014)	Dutch dataset		Beijing dataset	
				Peak	Off-peak	Peak	Off-peak
Adj- R^2	—	—	—	0.577	—	0.594	0.162
RMSE	4.04 s	16.6 s	—	6.2 s–8.8 s	8.49 s	7.21 s	6.21 s
MAPE	14.55%	—	15%	11.5%–14.2%	19.9%	11.4%	13.1%

Note. “Adj- R^2 ” denotes the adjusted coefficient of determination; “RMSE” denotes the root mean square error; “MAPE” denotes the mean absolute percentage error.

TABLE 9: The distribution of the worldwide subway dwell time.

City	Station	Date	Dwell time	Flow/dwell	Passenger	Average headway
San Francisco	Montgomery	pm peak	30 s–60 s	38%	586	153 s
New York	Grand Central	am peak	35 s–120 s	64%	1143	160 s
Toronto	Commission King	am peak	25 s–45 s	31%	428	168 s
British Columbia	SkyTrain Burrard	am peak	20 s–30 s	40%	562	151 s

of the samples in the dataset. In other words, the model can estimate dwell time very well in the Beijing urban rail transit system during peak hours. Second, the RMSE and MAPE are tolerable in both the Dutch dataset and the Beijing dataset during the investigated periods. Therefore, the errors between the predicted and observed dwell times are sufficiently small that the model can be used to estimate the dwell time for selected intermediate ordinary stations in the Dutch railway network and urban railway line 4 in Beijing. Third, adj- R^2 in the model based on the Beijing dataset during the off-peak period is low in the model based on the Beijing dataset, where the cases in Netherlands reveal that the characteristics of the off-peak passenger flow in the Beijing dataset cannot be reflected well by the model.

5. Generality Analysis Using a Theoretical Approach

The comparison approach in the previous section performed well with regard to the generality of the model. However, this approach may face problems such as how many cases should be covered, although it is usually impossible to cover all cases for a model in practice. To solve this problem, the following section shows that the generality of the model is tested using a theoretical approach. The theoretical approach includes four aspects. These aspects include variable analysis, parameter analysis, scenario analysis, and model structure and condition analysis.

5.1. Variable Analysis. First of all, the generality of the model is discussed based on the variables.

Assumption 1. Under the premise that the parameters are fixed, the range of the values of the variables (definition domain) should be within a certain trusted value domain. In other words, it has a lower bound and an upper bound.

Based on Assumption 1, we can discuss how the estimation result will change and in which range can the result be trusted. When the variables change within the domain of the definition, we can determine the range of the model’s

variables, establishing the applicability of the model. For the model used in this research, the variables include two types, the length of the train L_k^s and L_{k-1}^s and the train dwell time DT_{k-1}^s , DT_k^{s-1} , and DT_k^{s-2} .

For the length of the train, literature [10] shows that the range of the length is 2–12 cars in the majority of the urban rail transit. For the dwell time, previous research [10] presented the distribution of the worldwide rail transit dwell time, as shown in Table 9. In this table, the authors introduced the four major rail transit systems and the range of the dwell time was recorded from 20 s to 120 s. So, we considered dwell times within this range to be acceptable.

Next, we discuss the result of the model when the variables change within the domain. The domain variables are introduced into the model which is regressed by the Beijing datasets to obtain the results. Figure 7 shows the results and it reflects that the dwell time of the target train in the target station falls within the limited range (20–120 s). The results for the variables from the domain are acceptable. Therefore, the model is universally relevant to the Beijing datasets when the variables are in the domain. The results obtained from the model regressed by the Dutch datasets are shown in Figure 8; most of the results are acceptable when the variables are from the domain. Only a few parts extend beyond the range, but the model still has a certain general relevance.

5.2. Parameter Analysis. In this section, the value ranges of the parameters are discussed. The values of the parameters can be inferred from both the structure of the model and the domain of the definition of the variables.

At first, there are two contributors, and the results are for the dwell time in the model. If the trains’ length does not change in the modeling scenario, the value of the dwell time at the target station can be calculated approximately as a weighted average of the dwell times of the preceding train, at the previous station and the second previous station. Thus, we consider that the values of β_3 and β_4 which are the parameters of the contributors should lie in $[0, 1]$, which also ensures that the dwell time ranges of the different stations in a given railway are consistent.

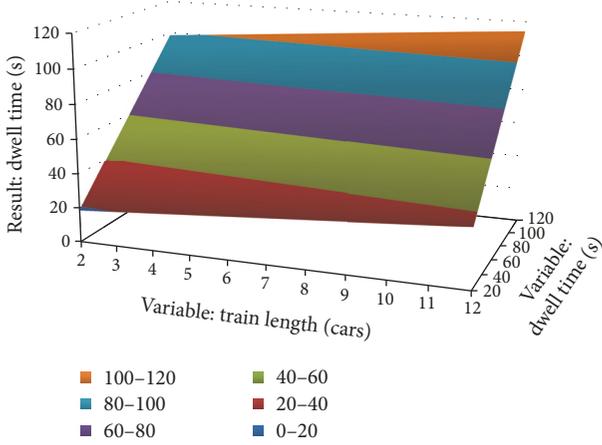


FIGURE 7: The results from changing the variables in the Beijing railway stations.

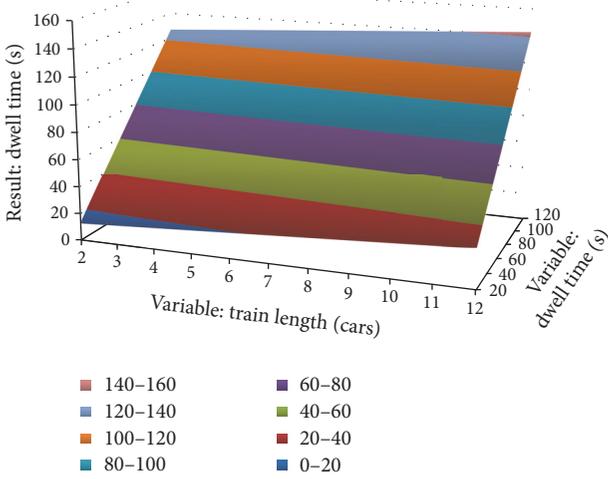


FIGURE 8: The results from changing the variables in the Dutch railway stations.

It can be deduced from the meaning of the parameter that the value of c is mainly related to the variance of passenger demand between the different stations and trains. Because this variation can be significantly different from station to station, the value can be deduced only by model regression. Formula (2) can be manipulated to obtain a new formula for parameter c , as shown in formula (4). In this formula, the dwell times of different trains at different stations are considered the same. In addition, the range of the dwell time is $[20, 120]$ and the range of the train lengths is $[2, 12]$; thus the range of parameter c is obtained as $[-220, 120]$. When the train lengths and the dwell time of other trains at other stations have a weak influence on the target, the dwell time is similar to parameter c . As a result, we limit the range of c in $[-20, 100]$.

$$c = DT_k^s - \beta_2 L_{k-1}^s - \beta_3 DT_{k-1}^s - \beta_4 \sqrt{DT_k^{s-1} * DT_k^{s-2}}. \quad (4)$$

Finally, the parameters concerning the train lengths are discussed. Although the ranges of the other parameters and

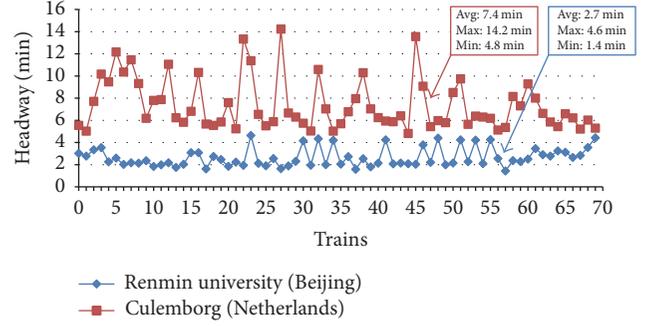


FIGURE 9: The headway of two cases.

variables are already limited, nonetheless we can still obtain the ranges of β_1 and β_2 . The ranges of these parameters are too large, and the relationships between train length and dwell time are not clear. The train length parameters are difficult to determine.

In summary, when the parameters are in their established ranges, the results can be trusted, and the ranges of the parameters are within the model scope of the application.

5.3. Scenario Analysis. The scenario is usually defined by several indicators, such as headway, type of station, and train speed. The scenario is one of the most important factors of influence when we discuss the generality of a model.

Assumption 2. The model is general if it can apply to more scenarios.

To study whether the model can be applied to different scenarios, it is important to discuss the effect of these scenario indicators on the model.

First, the effect of headway is twofold. On one hand, the headway influences the relationship between the dwell time of the target train at the target station DT_k^s and the dwell time of the previous train at the target station DT_{k-1}^s . The larger the headway, the weaker the relationship. The results of the model when regressed by the Beijing and Dutch datasets separately reflect this conclusion, and the headways of the two cases are shown in Figure 9. The red line denotes the headways of departing trains at the Culemborg station in the Dutch railway, and the blue line denotes the headways of departing trains at the Renmin University station in Beijing. The average headway of the Dutch dataset is 7.4 minutes, and DT_k^s and DT_{k-1}^s have a weak relationship; thus $\beta = 0.03$. The average headway of the Beijing dataset is 2.7 minutes, and the relationship between DT_k^s and DT_{k-1}^s is strong; thus $\beta = 0.76$ and the set DT_{k-1}^s has a strong impact on DT_k^s . This result can be explained by the fact that the relationship between passenger demand and consecutive trains is stronger when the headway is small. On the other hand, the minimum headway is determined based on the dwell time, operating margin, and safety separation time, as shown in formula (5). h_{\min} denotes the minimum headway, DT_k^s means the dwell time, T_{om} stands for the operating margin time, and T_{ss} identifies the safety separation time. The dwell time and

the minimum headway have a positive correlation, so the minimum headway has an influence on the dwell time.

$$h_{\min} = DT_k^s + T_{\text{om}} + T_{\text{ss}}. \quad (5)$$

However, the headways in Figure 9 from either Netherlands (during peak hours) or China are less than 15 min. For a larger headway, such as a headway greater than 15 min, formula (5) does not hold. The evidence can also be found in Li et al. (2016).

Second, the effect of different types of stations is also an indicator for determining which scenario is relevant. At intermediate stations, where a train cannot overtake another, the dwell time is mainly determined by the dwell time of the previous train, the minimum headway, and the passengers. At interchange stations, the previous train's dwell time is one of the most significant influencing factors. Moreover, whenever the train timetable is required for synchronization between two lines, the connections for transfer passengers should be considered. That is, the dwell time of the target train is related to the departure and arrival times of other trains in the interchange station. At the terminal station, the dwell time is mainly determined by the time when the train turns back. The turn back time is related to the train's length and turn back technique. The shorter the train and the higher its efficiency in turning back, the shorter the dwell time.

The average headway, minimum headways, train length, and dwell time (the target and previous train at the target and previous station) can all be used to describe different scenarios. The model is applicable to those scenarios, if the indicators that determine the scenarios are considered in the model. Based on our analysis, the model is applicable to more scenarios because the correlation between the headway and dwell time is high, and we can also determine from formula (5) that the dwell time is a component of the minimum headway. As a result, the headway is not considered in the regression model.

5.4. Structure and Condition Analysis. In the selected model, passenger demand is disregarded because passenger demand can be reflected by the substitute variables such as the dwell time of the preceding train and the dwell time at the previous station, so its assumption is not proven to be true. To demonstrate the generality of the model, it is necessary to know to what extent this assumption holds. To answer this question, the dwell time variables of the model can be replaced by an equation for the other variables that have already been proven to be relevant from formulations in previous research. Then, we can establish a new model with new variables and discuss the applicability of the new model. If the new model is plausible, then the assumption is proven.

In previous research, Puong [11] developed a nonlinear passenger regarded model to estimate train dwell time. The variables of the model included the number of standing passengers in the vehicle and their interactions with the boarding and alighting passengers, as shown in formula (6).

$$DT = 12.22 + 2.27B_d + 1.82A_d + 6.2 \cdot 10^{-4} \cdot TS_d^3 B_d, \quad (6)$$

where A_d represents the number of alighting passengers per door, B_d represents the number of boarding passengers per

door, and TS_d represents the number of standees per door. The model is used to estimate the dwell time at stations on the MBTA Red Line. The result shows that R^2 is approximately 0.9; the accuracy of the estimation is high.

We enter formula (5) into formula (2) and then obtain a new formula, as shown in formula (7).

$$\begin{aligned} \text{Left part} &= 12.22 + 2.27B_{d,k}^s + 1.82A_{d,k}^s + 6.2 \cdot 10^{-4} \\ &\quad \cdot TS_{d,k}^s{}^3 B_{d,k}^s, \\ \text{Right part} &= -0.83 + 1.6 \times L_{k-1}^s + 0.76 \times (12.22 \\ &\quad + 2.27B_{d,k-1}^s + 1.82A_{d,k-1}^s + 6.2 \cdot 10^{-4} \\ &\quad \cdot TS_{d,k-1}^s{}^3 B_{d,k-1}^s) + 0.04 \times [12.22 + 2.27B_{d,k}^{s-1} \\ &\quad + 1.82A_{d,k}^{s-1} + 6.2 \cdot 10^{-4} \cdot (TS_{d,k}^{s-1})^3 B_{d,k}^{s-1} * 12.22 \\ &\quad + 2.27B_{d,k}^{s-2} + 1.82A_{d,k}^{s-2} + 6.2 \cdot 10^{-4} \\ &\quad \cdot (TS_{d,k}^{s-2})^3 B_{d,k}^{s-2}]^{0.5}, \end{aligned} \quad (7)$$

where $B_{d,s}^k$, $A_{d,k}^s$, and $TS_{d,k}^s$ indicate the number of boarding passengers, alighting passengers, and standees per door of train k at station s .

In this formula, apparently $B_{d,s}^k$, $B_{d,s-1}^k$, $B_{d,s-2}^k$, and $B_{d,s}^{k-1}$ are different. The difference among $B_{d,s}^k$, $B_{d,s-1}^k$, and $B_{d,s-2}^k$ is the boarding passengers of the same train at different stations and is influenced by the passenger OD demand. The difference between $B_{d,s}^k$ and $B_{d,s}^{k-1}$ is the boarding passengers of different trains at the same station and is influenced by the period of time. We enter the value ranges of A_d , B_d , and TS_d in Puong [11] into formula (7), and the result is shown in Figure 10. It can be seen that the value of the left part and the value of the right part are close in half of the cases, whereas the value of the left part of the formula is not equal to the value of the right in the remaining cases. By applying a two-sample t -test, it can be shown that $H = 0$ means that the two samples are independent random samples from normal distributions with equal means at a 5% significance level, demonstrating that the model is reliable.

6. Generality Improvement

From the generality testing, it can be seen that the relationships between passenger demand at different stations and trains significantly impact the estimation result. However, such a relationship is not necessarily based on consecutive stations or consecutive trains. Based on this inference, we introduce a more generalized predictor to improve the generality of the proposed model, and the results of the improved model are compared with the results of the proposed model in this section.

6.1. A More Generalized Predictor. The generality of our model can still be improved. Past research analyzed the relationships among the dwell times of three consecutive stations

TABLE 10: Correlation coefficients of the dwell time in peak periods.

Stations	Pervious stations										
	1	2	3	4	5	6	7	8	9	10	11
Renmin University	0.22	0.47	0.44	0.16	0.12	—	—	—	—	—	—
Weigongcun	0.22	0.42	0.29	0.27	0.41	0.77	—	—	—	—	—
National Library	0.06	0.36	0.05	0.61	-0.12	0.33	0.54	—	—	—	—
Beijing Zoo	0.02	-0.06	-0.21	0.31	-0.13	0.04	0.16	0.75	—	—	—
Xizhimen	-0.07	-0.03	-0.15	0.16	-0.15	-0.13	0.12	0.44	0.49	—	—
Xinjiekou	-0.03	-0.13	-0.23	0.14	-0.04	0.00	0.08	0.28	0.59	0.71	—
Ping'anli	-0.15	-0.14	-0.15	-0.08	0.34	-0.09	0.02	0.07	0.24	0.33	0.53

Notes. Numbers 1–11 denote stations on line 4 from Xiyuan to Xinjiekou.

TABLE 11: Correlation coefficients of the dwell time in off-peak periods.

Stations	Pervious stations										
	1	2	3	4	5	6	7	8	9	10	11
Renmin University	0.33	0.12	0.38	0.06	0.12	—	—	—	—	—	—
Weigongcun	0.31	0.06	0.14	0.12	0.06	0.23	—	—	—	—	—
National Library	-0.11	0.21	0.1	0.41	0.15	0.22	0.12	—	—	—	—
Beijing Zoo	-0.12	0.03	-0.14	0.08	-0.15	-0.05	0.64	0.03	—	—	—
Xizhimen	0.32	-0.03	-0.1	0.27	0.13	0.09	-0.24	0.1	0.2	—	—
Xinjiekou	-0.06	0.18	0.22	0.39	0.18	0.17	0.07	0.24	0.28	0.46	—
Ping'anli	0.08	0.1	0.13	0.15	0.1	0.16	0.09	0.12	0.08	0.1	0.17

Notes. Numbers 1–11 denote stations on line 4 from Xiyuan to Xinjiekou.

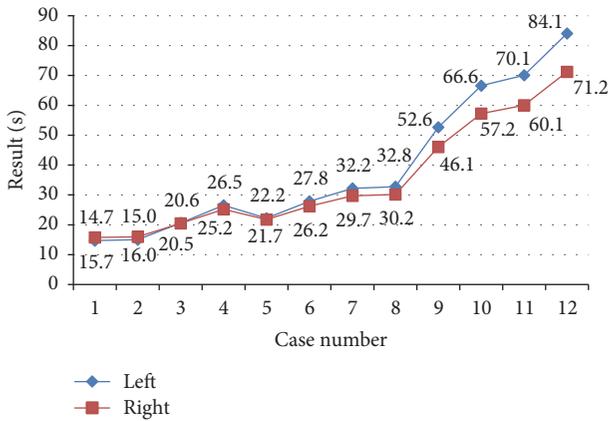


FIGURE 10: The comparison of the values of the left part and right part in formula (7).

in the Dutch railway lines. Because of the data consistency problem, some data were missing due to incorrect track identifiers resulting from track changes after maintenance. However, the Beijing dataset includes the dwell times of all the stations on the selected line, which inspires our refinement of the model.

We can analyze the relationships between the dwell times of the target station and previous stations. The correlation of dwell time between the target station and previous stations in peak and off-peak periods is analyzed and the result is shown in Tables 10 and 11, respectively. Through correlation analysis, we can find the most correlated previous station,

which is called the associated station (the boldface number in the table). The dwell time of the associated station (denoted as DT_k^{s*}) impacts the target station more than the square root of the product of the dwell time at the previous station and at the second previous station on ($\sqrt{DT_k^{s-1} * DT_k^{s-2}}$), which is used in the original model.

To test the above idea, the correlation coefficients between the dwell time at the target station and the dwell time at the associated stations are obtained and are compared with the square root term used in the original model. The results from the Beijing dataset are shown in Table 12.

Table 12 shows that the variable DT_k^{s*} has a stronger relationship with the dwell times of the target train at the target station than the original square root term. Thus, the predictor DT_k^{s*} is a better substitute for the predictor $\sqrt{DT_k^{s-1} * DT_k^{s-2}}$. The boldface indicates that there was a significant correlation (>0.3) between the predictors and the dwell time.

Based on this improvement, a more generalized model is obtained, as shown in the following formula:

$$DT_k^s = c + \beta_1 L_k^s + \beta_2 L_{k-1}^s + \beta_3 DT_{k-1}^s + \beta_4 DT_k^{s*}. \quad (8)$$

The Beijing dataset is used to calibrate formula (8), and the results from the peak and off-peak periods are shown in formula (9) and formula (10), respectively.

$$DT_k^s = -8.04 + 1.6 \times L_{k-1}^s + 0.61 \times DT_{k-1}^s + 0.34 \times DT_k^{s*}, \quad (9)$$

$$DT_k^s = 0.17 + 1.6 \times L_{k-1}^s + 0.47 \times DT_{k-1}^s + 0.26 \times DT_k^{s*}. \quad (10)$$

TABLE 12: Correlation coefficients of the predictors.

Stations	Peak periods			Off-peak periods		
	DT_{k-1}^s	$\sqrt{DT_k^{s-1} * DT_k^{s-2}}$	DT_k^{s*}	DT_{k-1}^s	$\sqrt{DT_k^{s-1} * DT_k^{s-2}}$	DT_k^{s*}
Renmin University	0.20	0.14	0.47	0.12	0.08	0.38
Weigongcun	0.41	0.56	0.77	0.15	0.12	0.31
National Library	0.54	0.42	0.61	0.58	0.16	0.41
Beijing Zoo	0.46	0.35	0.75	0.07	0.15	0.64
Xizhimen	0.47	0.47	0.49	0.23	0.14	0.32
Xinjiekou	0.73	0.65	0.71	0.02	0.36	0.46
Ping'anli	0.28	0.42	0.53	0.07	0.13	0.17
All stations	0.87	0.42	0.74	0.37	0.27	0.46

TABLE 13: Comparison with the improved parametric model.

Dataset	Dutch			Beijing			
	PM	NPM		PM	IPM		
Period	Peak	Off-peak		Peak	Off-peak	Peak	Off-peak
Indicators							
Adj- R^2	0.577	—	0.594	0.162	0.860	0.455	
RMSE	6.2 s–8.8 s	8.49 s	7.21 s	6.21 s	4.19 s	4.14 s	
MAPE	11.5%–14.2%	19.9%	11.4%	13.1%	7.8%	9.9%	

6.2. *Estimation Performance.* To test the performance of the improved model and compare the estimation accuracy of the improved model with that of the original model, adj- R^2 , RMSE, and MAPE indicators are used. The accuracy of the improved parametric model (IPM) is compared with that of the parametric model (PM) based on the Beijing dataset for the peak and off-peak periods (see Table 13).

In the IPM, adj- R^2 is higher than that in the PM during both the peak and off-peak hours. Thus, the IPM can reflect more characteristics of the Beijing dataset. However, adj- R^2 is still low in the IPM during the off-peak periods, and improving the accuracy of the off-peak period by adding new influencing factors is difficult. Other prediction methods for the off-peak hours may require testing. Moreover, the RMSE and MAPE of the IPM are smaller than those of the original model, indicating that there are smaller differences between the estimated and actual dwell times. Therefore, the results of the IPM are better than the results of the PM. In other words, using the IPM improves the accuracy of the results.

Due to the data consistency problem in the Dutch railway, the associated stations are not analyzed further. The square root term can be a special case of the associated station; therefore, it can be deduced that using the associated station item would not reduce the model's accuracy.

7. Discussion

7.1. *Scenario of the Model.* Based on the test datasets in this paper, it is worth noting that the generality of the estimation model indicates that the model with the same parameters and predictors can fit for all railway lines, regardless of whether the line is a subway or commuter rail or other railway types. Certain conditions must be satisfied to use the model.

The original model is specifically designed for short stops. Short stops are stops on the open track where sidings are usually not available and where trains dwell only for alighting and boarding, after which they immediately continue their journey. In both cases, all stations are short stops without sidings, and whether the model fits stations with sidings and large stations with passenger connections remains unclear.

The estimation model with the same predictors can regress the dwell time of different stations on different lines even in different countries, and the result is valid. However, the values can be different in different cases. In other words, the model followed the specific cases, and the parameters vary from case to case. The main reason is that the influence level of the predictors may differ in different cases. Therefore, when applying the model to a new case, the parameters should still be regressed. Accordingly, it is better to use the model at rescheduling phase where the train system is running and necessary datasets such as previous dwell times and train length are available.

In summary, the approach in this paper is convincing. However, it is difficult to establish a model that can fit all scenarios.

7.2. *Dataset Requirement of the Model.* The dataset should follow several constraints when studying specific cases. To improve the model's performance, the following principles should be considered.

The first issue concerns the data investigation site:

(1) The original model focused on short stops where sidings are not usually available and trains dwell for only alighting and boarding. That is, no overtaking occurs in the station. Therefore, the selected stations should be small stations with the same characteristics.

(2) In the original case study, the passengers check in, wait, and board the train without barriers. Therefore, the new dataset should have the same system.

(3) Three consecutive stations were selected in the original model to examine the relationship between the dwell times of consecutive stations; therefore, the new dataset should contain the dwell times from at least three stations.

(4) In the original case, the rolling stock is sprinters with wider doors and small gap differences between the rolling stock and the station to ensure fast alighting and boarding processes. These trains stop at every station without skipping any stops; therefore, the new case of the rolling stock should also be used for the rapid boarding and alighting of passengers.

Another issue is measurement. The measurements in the two cases are different, which could cause bias. The dwell time is estimated from track occupation data in the Dutch case; in that case, the train dwell time is estimated by an algorithm according to the stop sign of a train and the time-dependent occupation states of a track [21]. The exact position of a train is not clear. In fact, it is not possible for all drivers to stop exactly in front of the stop sign. From this point of view, the automatic measurements may underestimate the dwell time. In the Chinese case, the dwell times are measured manually via stop watches. The time is recorded as soon as a train standstill or departure occurs. Despite the error from the investigators, the accuracy could be higher than in the Dutch case. Based on the analysis, the error distributions of the two cases would be significantly influenced by the measurement. The theoretical issues associated with using automated versus manual dwell time measurements should be considered in future research.

7.3. Restrictions of the Model. Station dwell times are the major component of headways at close train frequencies. The existing literature suggests that the best achievable headways under these circumstances are in the range of 110 to 125 seconds [10]. Thus, it is reasonable that, as the major part of the headway, the dwell time also lies within a loosely identified interval. We discuss the ranges of train lengths, dwell times, and headways under variable analysis in Section 5.1, according to a widely selected dataset from the references [10]. In addition, we found that the model is applicable within this range. However, it should be noted that this inference does not necessarily hold at low railway frequencies, especially for long distance railways. Because the headways in the two datasets are less 15 min, we can only conclude that the model is applicable for close frequencies. Furthermore, we can also obtain the evidence from former research [2] to prove that formula (5) may not hold in larger headway scenarios (larger than 15 minutes).

8. Conclusions

The main contribution of this paper is to propose a systematic approach for testing the generality of a train dwell time estimation model and improving the model performance by introducing a more generalized predictor.

The generality of the model is tested by using two approaches, namely, a comparison approach and a theoretical

approach. In the comparison approach, a dwell time estimation is selected and applied to a completely new independent test bed. The test results from two datasets from two different countries show relatively high accuracy and equivalent effectiveness. The regression parameters are compared using the dataset of line four in the Beijing urban rail transit system. The parameters are regressed using the datasets of these two cases. In the theoretical approach, the train dwell time estimation model is analyzed by four steps: variable analysis, parameter analysis, scenario analysis, and model structure and condition analysis. We conclude that the test model is general in the given condition, namely, short stops with a headway less than 15 min.

Furthermore, a more generalized predictor, dwell times at the associated station, is introduced to improve the performance of the model. The performance indicators show that the result of the improved model is better than the result of the original model. Thus, the associated station would be relevant for estimating train delays during operation, and using the associated station item would improve the generality of the results.

Usually, generality and accuracy are two contrary indicators for a model. If the accuracy of the model is higher, the range of the model's applicability is narrower, meaning that the model can be applicable only in certain scenarios. In contrast, if the model is applicable to many different scenarios, in other words, it is more general, it is difficult to maintain high accuracy at the same time as generality. The estimation error of the model we test is acceptable. Meanwhile, it is proven that the model is relatively general; therefore, the model can balance the tradeoff between generality and accuracy well.

The focus of future research should be on two directions. On one hand, more datasets from different cases can be tested to verify the model. On the other hand, the accuracy and the generality of the model should be improved, particularly for long headways. Randomness and the weak relationship between the dwell time and passenger demand under this condition make dwell time estimation more difficult. To improve the accuracy of the model, other nonparameter methods, such as kernel regression and the nearest neighbors method, which do not require a significant amount of data, can be tested. More influencing factors, such as the transfer rule, weather, and accident, could also be considered if such data can be collected by using sensors. These factors will be extensively studied in the near future.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (U1434207), the Fundamental Research Funds for the Central Universities of China (2016JBM030), the National Key Research and Development Plan (2016YFE0201700), and the Beijing Chaoyang District Science and Technology Commission (CYXC1607). The authors wish to acknowledge these agencies.

References

- [1] C. T. Lewis and C. Belanger, "The generality of scientific models: a measure theoretic approach," *Synthese. An International Journal for Epistemology, Methodology and Philosophy of Science*, vol. 192, no. 1, pp. 269–285, 2015.
- [2] D. Li, W. Daamen, and R. M. P. Goverde, "Estimation of train dwell time at short stops based on track occupation event data: A study at a Dutch railway station," *Journal of Advanced Transportation*, vol. 50, no. 5, pp. 877–896, 2016.
- [3] Q. Zhang, B. Han, and D. Li, "Modeling and simulation of passenger alighting and boarding movement in Beijing metro stations," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 5, pp. 635–649, 2008.
- [4] S. Seriani and R. Fernandez, "Pedestrian traffic management of boarding and alighting in metro stations," *Transportation Research Part C: Emerging Technologies*, vol. 53, pp. 76–92, 2015.
- [5] A. Yamamura, M. Koresawa, T. Inagi, and N. Tomii, "Dwell time analysis in railway lines using multi agent simulation," in *Proceedings of the 13th World Conference on Transportation Research (WCTR)*, Rio de Janeiro, Brazil, 2013.
- [6] F. Sourd, C. Talotte, C. Ambroise et al., "Modelling of pedestrian flows during dwelling: development of a simulator to evaluate rolling stock and platform flow performance[C]//WCRR2011-World Congress on Railway Research, Lille France. 2011.
- [7] W. H. K. Lam, C. Y. Cheung, and Y. F. Poon, "A study of train dwelling time at the hong kong mass transit railway system," *Journal of Advanced Transportation*, vol. 32, no. 3, pp. 285–296, 1998.
- [8] S. C. Wirasinghe and D. Szplett, "An investigation of Passenger interchange and train standing time at LRT stations: (II) estimation of standing time," *Journal of Advanced Transportation*, vol. 18, no. 1, pp. 13–24, 1984.
- [9] T. M. Lin and N. H. M. Wilson, "Dwell time relationships for light rail systems, 1992.
- [10] T. Parkinson and I. Fisher, "Rail transit capacity," TCRP Report, 13, Transportation Research Board, Washington, DC.
- [11] A. Puong, *Dwell Time Model and Analysis for the mbta Red Line*, Massachusetts Institute of Technology Research Memo, Cambridge, UK, 2000.
- [12] N. G. Harris, "Train boarding and alighting rates at high passenger loads," *Journal of Advanced Transportation*, vol. 40, no. 3, pp. 249–263, 2006.
- [13] J. C. Jong and E. F. Chang, "Investigation and estimation of train dwell time for timetable planning," *Office of Scientific & Technical Information Technical Reports*, vol. 4141, pp. 178–185, 2011.
- [14] X. Karekla and N. Tyler, "Reduced dwell times resulting from train-platform improvements: The costs and benefits of improving passenger accessibility to metro trains," *Transportation Planning and Technology*, vol. 35, no. 5, pp. 525–543, 2012.
- [15] W. Daamen, Y.-C. Lee, and P. Wiggeraad, "Boarding and alighting experiments: Overview of setup and performance and some preliminary results," *Transportation Research Record*, no. 2042, pp. 71–81, 2008.
- [16] S. Buchmueller, U. Weidmann, and A. Nash, "Development of a dwell time calculation model for timetable planning," in *Proceedings of the 11th International Conference on Computer System Design and Operation in the Railway and Other Transit Systems, COMPRAIL 2008*, pp. 525–534, Spain, September 2008.
- [17] F. S. Hänseler, M. Bierlaire, and R. Scarinci, "Assessing the usage and level-of-service of pedestrian facilities in train stations: a swiss case study," *Transportation Research Part A: Policy and Practice*, vol. 89, pp. 106–123, 2016.
- [18] I. A. Hansen, R. M. P. Goverde, and D. J. Van Der Meer, "Online train delay recognition and running time prediction," in *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems, ITSC 2010*, pp. 1783–1788, Portugal, September 2010.
- [19] P. Kecman and R. M. P. Goverde, "Predictive modelling of running and dwell times in railway traffic," *Public Transport*, vol. 7, no. 3, pp. 295–319, 2015.
- [20] P. Kecman and R. M. P. Goverde, "Process mining of train describer event data and automatic conflict identification," in *Computers in Railways XIII, WIT Transactions on the Built Environment*, C. A. Brebbia, N. Tomii, P. Tzieropoulos, and J. M. Mera, Eds., vol. 127, pp. 227–238, 2012.
- [21] D. Li, R. M. P. Goverde, W. Daamen, and H. He, "Train dwell time distributions at short stop stations," in *Proceedings of the 2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014*, pp. 2410–2415, China, October 2014.



Hindawi

Submit your manuscripts at
www.hindawi.com

