

Modelling Passenger Distribution on Metro Platforms Based on Passengers' Choices for Boarding Cars

Jiping Fang

Centre for Transport Studies, Department of Civil, Environmental & Geomatic Engineering, University College London, London, UK

jiping.fang.15@ucl.ac.uk

Taku Fujiyama

Centre for Transport Studies, Department of Civil, Environmental & Geomatic Engineering, University College London, London, UK

taku.fujiyama@ucl.ac.uk

Howard Wong

*Centre for Advanced Spatial Analysis, University College London, London, UK;
Public Transport Service Planning, Transport for London, London, UK*

howard.wong@tube.tfl.gov.uk

Modelling Passenger Distribution on Metro Platforms Based on Passengers' Choices for Boarding Cars

Metro capacity should be further improved with growing demand for metros. For metro lines with high-frequency services, the train dwell time is a determinant of the number of trains per hour. Uneven passenger distribution on the platform may cause a longer dwell time, thereby reducing the capacity. This study analyses the mechanism of the passenger distribution on the platform with regard to which cars of a train passengers choose to board. A model has been developed, considering the origin-destination and waiting time of passengers and the platform layouts of the stations where passengers board and alight. The model was calibrated using loadweigh data, which provides the weight of passengers on each car of a train, on the Hammersmith & City line of the London Underground. The largest difference between the actual and modelled average number of boarders on a car is less than 2 passengers, which suggests the model performs well. The results show, at both stations studied, 44% of passengers in the morning peak chose boarding cars based on the layouts of the stations where they alighted. The results would be useful for metro operation planners and station staff to take measures to manage passenger distribution on the platform.

Keywords: metro platform, passenger distribution, boarding behaviours, loadweigh

1. Introduction

The demand for metros is growing in many metropolises. According to Transport for London (2018), the total passenger journeys on the London Underground have increased by 10.6% over the last 6 years (from 1.23 to 1.36 billion). Meanwhile, in the Beijing Subway in China, the number rose by 53.7%, from 2.46 to 3.78 billion (BMBS 2018). There is a global need to further improve the metro network capacity.

The train dwell time is the elapsed time from the instant a train stops moving at a station until the instant it resumes moving. The dwell time has three components: passenger alighting and boarding time, door opening and closing time, and safety check

time (Wirasinghe and Szplett 1984; Bian et al. 2015). The alighting and boarding time is the predominant element and is usually determined by the numbers of passengers boarding, alighting, staying inside the train, and staying on the platform (Fujiyama, Nowers, and Tyler 2008), unless the train is held for railway traffic management reasons. Because a train has multiple doors and the alighting/boarding time of a train is decided by that of the door with the longest alighting and boarding time (unless alighting and boarding finishes at all the doors, but the doors cannot be closed), the distribution of passengers along the cars of a train affects the dwell time. Moreover, severe uneven passenger distribution not only results in high crowding levels in particular areas on the platform and creates difficulties for passengers to walk to the less crowded areas, but it could also lead to potential safety risks. Therefore, managing passenger distribution on the metro platform could be an alternative approach to reducing dwell times.

Some research on passenger distribution on the metro platform has been carried out previously, using field counting as a typical practice to quantify and analyse it. The count results given in various literature (Szplett and Wirasinghe 1984; Lam, Cheung, and Lam 1999; Wu, Yang, and Yuan 2010; Krstanoski 2014) showed that there was a negative correlation between the congestion level of a waiting area on the platform and the walking distance to that area from the platform entrance, the most crowded waiting area being the one adjacent to the entrance. However, some cases did not follow this general rule. For example, Lam, Cheung and Lam (1999) analysed the passenger concentration on a platform of the Hong Kong Light Rail Transit in the morning peak time by dividing the platform into six blocks. It was observed that the most crowded block was the second-closest one to the platform entrance, not the closest one. Moreover, Szplett and Wirasinghe (1984) observed that, with only one entrance in the

station of a light rail transit line in Calgary, two boarder clusters existed on the platform. The underlying mechanism of the latter two cases was unclear.

Modelling is another tool for studying passenger distribution on the metro platform. Lu and Dong (2010), Wu, Yang and Yuan (2010), Wu et al. (2012) and Krstanoski (2014) developed models that predicted the probability of a passenger choosing a certain boarding door of a train, directly based on the field counting data of the number of passengers boarding through each door. They suggest that the passengers' choices of boarding doors depends on the walking distance from the platform entrances to the waiting area and the level of crowdedness around doors at the station where passengers boarded the trains.

However, the destinations of passengers and the layouts of the station at which they would alight from the trains were rarely considered in existing research regarding passenger distribution on the metro platform. In fact, according to the field survey of Kim et al. (2014), 53.4% of the interviewees stated that they chose boarding cars in order to minimise the walking distance to the exits at the stations where they alighted from the trains, while only 12.7% stated they preferred shortening the walking distance at the stations where they boarded the trains. Kim et al. (2014) also found that many passengers prefer to minimise the walking distance at the destination station.

This paper is structured in the following way: Section 2 introduces the boarding strategies of passengers. The datasets providing the Origin-Destination (OD) pairs of passengers and the number of passengers on each car of each train are presented in Section 3, together with a survey of the platform layouts at the stations on the Hammersmith & City line. The methodology of two sub-models is explained in Section 4, and the calibration and evaluation methodology of the model is in Section 5. Section

6 shows the model performance and results, and the results are discussed in Section 7. Section 8 draws conclusions.

2. The boarding strategies of passengers considered in this study

According to Kim et al. (2014), the main driving factors for the choice of a boarding car are to minimise the walking distance and to maximize safety and comfort. Some passengers may prefer to minimise their walking distance on the platform at the station where they board the train. At the same time, they look for a less crowded area to stand and board. This strategy is called the origin-based choice.

On the other hand, some other passengers may choose boarding cars closest to the platform exits or interchange pathways at the station where they alight from the train, so that they can minimise their walking distance at that station. This strategy is called the destination-based choice. Commuters who frequently travel in the metro system may use this strategy, because they are familiar with the layouts of both their origin and destination stations.

This study assumes the above two strategies to be key in passenger choice of boarding cars. The study has developed two sub-models of the number of boarders and alighters on each car of each train for these two boarding strategies. The passenger distribution on the metro platform can be predicted by combining these two sub-models. In addition, passengers who are familiar with the layouts of the stations may act differently from those who are not, because they do not need to refer to the metro map on the platform, and they may choose boarding cars that can minimise the walking distance on the platform at the station where they will alight to save journey time.

It should be noted that there may be many other variables that influence passenger choice of boarding cars. However, the consideration of many factors in the modelling can lead to overwhelmingly complex models. In addition, we require our

models to use data which is readily available (e.g. station asset data, train crowding). For example, some passengers with high heels may dislike boarding specific cars because of the wide gap between the car and the platform, but there is no data about such personal characteristics (i.e. how many people wear high heels). Thus, as the first step of research into boarding car choice, this research focuses on the aforementioned variables only: the layouts of the platforms at the boarding and alighting stations of passengers.

3. The case study and datasets

3.1. The case study line and stations

Passengers using the Hammersmith & City line of London Underground (the H&C line) in the morning peak hours were selected as the case study for this study. The H&C line is 25.5 km long with 29 stations connecting Hammersmith in West London to Barking in East London. From Hammersmith Station to Liverpool Street Station, the H&C line shares the same track with the Circle line, and Hammersmith Station is the terminus of both lines. During peak hours, a train departs from Hammersmith Station every 5 minutes, and the trains of the two lines run on the track alternately.

The H&C line runs seven-car trains. There are 3 passenger doors on each car, 21 doors in total on an S stock train. There are only two relatively narrow doors (width: 1210 mm) on the train, which are installed on the two front cars. The width of the wider doors is 1610 mm. Thus, the influence of door width on passenger distribution can be ignored. The distance is 107.85 m between the middle line of the first door on Car 1 and that of the last door on Car 7. A schematic plot of a train at an H&C station is shown in Figure 1.

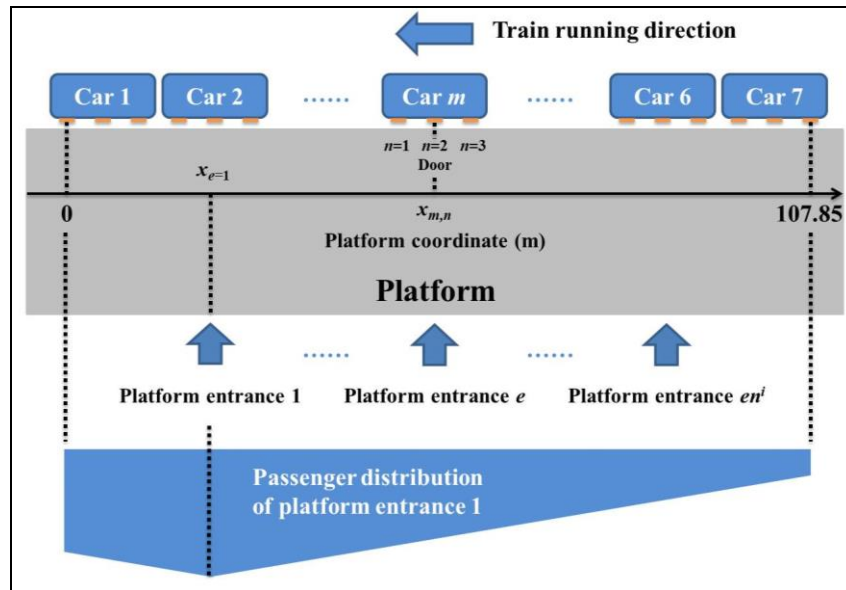


Figure 1: Diagram of a train on the H&C line and the platform.

In this study, the station on the H&C line where a passenger boards an H&C line train is called the ‘origin station’, regardless of whether they start their journey there, or have interchanged from another metro line. For example, if a passenger starts his/her journey at Green Park Station on the Jubilee line, interchanges to the H&C line at Baker Street Station and travels to Euston Square Station, his/her origin station is Baker Street Station in this study. Similarly, a station on the H&C line where a passenger alights from the H&C line train is called his/her destination station, regardless of whether he/she completes his/her journey at the station or interchanges to another line from there.

The eastbound platforms at Hammersmith Station and Goldhawk Road Station were selected to test the model performance in this study. Hammersmith is a terminal station on the H&C line, and the closest car of a train to its only platform entrance is Car 7. Goldhawk Road is the second station when an H&C line train runs from Hammersmith to Barking and also has only one platform entrance. However, the closest car to the platform entrance at Goldhawk Road is Car 1. There are no screen doors and pillars on the platform at these two stations. At both stations, passengers usually can

find seats on the trains and the influence of passenger loading on board is quite limited on passengers' choices for boarding cars.

3.2. Station layout observation

An on-site survey of platform layouts was conducted for all the stations on the H&C line. Relevant information, including the closest cars to the platform entrances, exits and interchange pathways, was collected. This survey can help understand how the passengers who are familiar with the layouts of their origin and destination stations choose their boarding cars based on the destination-based choice.

3.3. Rolling origin destination survey

Rolling Origin Destination Survey (RODS) data is used to provide the OD pairs among the stations on the H&C line in this study. It was commissioned by Transport for London to collect information on how and why passengers travelled around the London Underground network. The fieldwork was conducted annually from 1998 between 07:00-24:00 on weekdays in October and November by questionnaires given to passengers. Passenger journey information was collected, such as passengers' OD pairs, journey purposes, actual route choice (interchange stations), and flow volumes through platform entrances.

This study uses the RODS that was conducted in October and November 2016, in order to calculate the percentages of the passengers travelling from a certain station to other stations along the H&C line.

In an individual train, these percentages are random. However, if the size of the train samples is large enough, the percentages can be calculated based on the statistical results of the RODS. Combined with the survey results of the platform layouts, a sub-model that describes the destination-based choice can be developed.

3.4. Loadweigh

The rolling stock (S stock) used on the H&C line can weigh the passengers in each car by an airbag-based suspension system, which is designed to provide an accurate train weight for modifying the acceleration and braking forces. The data is called loadweigh. It is collected every time the train departs from a certain station.

Divided by the average weight of passengers, loadweigh can be converted into an estimated number of passengers on each car of an S stock train at every link on the H&C line. According to the Health Survey for England 2015, which was commissioned by National Health Service Digital in the UK, the average weight of adults in England was 78.1 kg (NHS Digital 2016). Given that children also travel on the metro, the average weight of passengers used in the model of this study is 75 kg, which number is also used by Transport for London to estimate the number of passengers on board.

Although loadweigh cannot directly show the actual number of boarders and alighters, it reveals the difference between the numbers of passengers on a car before and after the train stops at a certain station. It should be noted that the passenger loading change is caused by passengers boarding and alighting, as well as passengers who walk to different cars while a train is running between stations. However, in the morning peak hours, the passenger loading in the train is usually quite heavy which makes it difficult for passengers to walk to another car. Thus, the influence of passenger on-train movement on passenger loading change can be ignored. Consequently, passengers alight from the same car where they board, and loadweigh can provide the difference between the numbers of boarders and alighters when the train stops at a certain station, which is used to calibrate the model of the boarders and alighters on each car of a train. In this study, the loadweigh of 342 trains that ran from Hammersmith to Barking on the

H&C line in the morning peak hours during September to December in 2016 is used to calibrate the model.

4. The model of the number of boarders and alighters on each car of a train

Passengers using the origin-based choice for boarding cars act differently from those using the destination-based choice, thus, for each boarding strategy, we developed a sub-model of the number of boarders and alighters on each car of a train. The number of boarders on each car of a train can denote the passenger distribution at a certain station.

4.1. The origin-based choice sub-model of the number of boarders on each car of a train

The walking distance, the waiting time and the number of passengers on the platform are the key elements affecting passenger behaviour, assuming that the individual characteristics of passengers are the same (e.g. gender and age). Consequently, these elements influence the passenger distribution on the platform. Each of the following pairs of elements leads to passenger clustering near the platform entrance.

4.1.1. The effect of the walking distance and the waiting time

The waiting time is limited. In order to ensure they can board the next train, passengers may decrease their walking distance as the waiting time reduces before the next train stops. As the waiting time elapses, the waiting area around the platform entrance becomes more crowded than the areas away from the entrance.

4.1.2. The effect of the walking distance and the number of passengers on platform

With an increase in the number of passengers on the platform, the walking space decreases, leading to difficulties for passengers to move along the platform (Lam,

Cheung, and Lam, 1999), which in turn causes a passenger cluster near the platform entrance.

4.1.3. The effect of the number of passengers on the platform and the waiting time

After a train has just departed, a passenger has sufficient time (until the next train) to move to a waiting area with fewer passengers. He or she can then be the first boarder onto a car in the next train. Thus, the number of passengers standing in front of him or her that is considered acceptable may be low. However, when a train is about to arrive, passengers usually stop walking and board the train at the doors nearest to them, which can result in the acceptability of a large number of passengers already standing in front of them. Hence the passenger distribution on the platform is likely to show a cluster close to the platform entrance when a train stops, and there can be a positive correlation between the number of passengers in a waiting area and the distance the area is away from the platform entrance.

It is assumed that the distribution of passengers using the origin-based choice is linear on the metro platform. The number of boarders through a door is inversely proportional to the distance between the middle line of the door and that of the platform entrance, as shown in Figure 1. The origin-based choice sub-model of the number of boarders on each car of a train is shown as follows:

$$Bo^i = \alpha B^i \quad (1)$$

$$\sum_{m=1}^7 (Bo_m^i)_e = q_e^i Bo^i \quad (2)$$

$$(Bo_m^i)_e = \sum_{n=1}^3 (Bo_{m,n}^i)_e \quad (3)$$

$$(Bo_{m,n}^i)_e = a_e^i |x_{m,n} - x_e| + b_e^i \quad (4)$$

$$Bo_m^i = \sum_{e=1}^{en^i} (Bo_m^i)_e \quad (5)$$

where

i, j, k are station numbers of the H&C line; for example, the number for Hammersmith is 1 and that for Barking is 29, $i, j, k \in [1, 29]$;

α is the percentage of passengers who choose their boarding cars using the origin-based choice, $\alpha \in [0\%, 100\%]$;

m is the car number, $m \in [1, 7]$. Car 1 is the first car that reaches a metro platform when the train is about to stop, as shown in Figure 1;

n is the door number, $n \in [1, 3]$, as shown in Figure 1;

en^i is the total number of the platform entrances at station i ;

e is the platform entrance number, $e \in [1, en^i]$, as shown in Figure 1;

q_e^i is the percentage of passengers who reach the platform through entrance e at station i , $\sum_{e=1}^{en^i} q_e^i = 1$. q_e^i can be calculated based on RODS;

B^i is the total number of boarders of a train at station i ;

Bo^i is the total number of boarders of a train at station i , who use the origin-based choice for boarding cars;

Bo_m^i is the number of boarders of a train at station i , who board car m using the origin-based choice;

$(Bo_m^i)_e$ is the number of boarders of a train at station i , who reach the platform through entrance e and board car m using the origin-based choice;

$(Bo_{m,n}^i)_e$ is the number of boarders of a train at station i , who reach the platform through entrance e and board on car m through door n using the origin-based choice;

$x_{m,n}$ is the coordinate of the middle line of door n of car m , $x_{m,n} \in [0, 107.85]$, unit: m;

x_e is the coordinate of the middle line of the platform entrance e , $x_e \in [0, 107.85]$, unit: m;

a_e^i and b_e^i are the parameters that describe the linear distribution of passengers reaching the platform through platform entrance e .

4.2. The ideal destination-based choice sub-model of the number of boarders on each car of a train

In order to ensure the shortest walking distance on the platform at the destination station, passengers need to consider not only the locations of the platform entrances at the origin stations, but also the locations of the platform exits linking to interchange pathways or station gates at the destination stations.

4.2.1. The influence of the layout of platform entrances and exits

Passengers enter the platform through platform entrances and leave it through platform exits. The layouts of platform entrances affect passengers' choices of their boarding cars. At some stations, there are more than one platform entrances and exits. For example, at King's Cross Station, there are two platform entrances on the westbound platform, which are located near Car 2 and Car 5 of a train going to Hammersmith Station, whereas Hammersmith Station has only one platform exit, and the nearest car to it is Car 1. If a passenger travels from King's Cross to Hammersmith and hopes to shorten the walking distance at the destination station, Car 2 is the target car.

4.2.2. The influence of the location of the interchange pathways

Two passengers may walk to different target cars even if they travel from the same origin station to the same destination station on the H&C line. This is because the locations of the platform exit linking to the station gates and that linking to the interchange pathway are different at the destination station, as at King's Cross Station. It can be noted that one of the passengers will interchange to another line at the destination station and the other will directly leave London Underground at that station.

Passengers on the platform on the H&C line can be classified into four types based on whether they interchange or not, because each type of passenger may choose different target cars at the origin stations. 1) The interchange-out passengers: the passengers start their journeys from the H&C line and interchange to other lines at their destination stations. 2) The interchange-in passengers: the passengers start their journeys from another line and interchange to the H&C line. 3) The interchange-in-out passengers: the passengers interchange to the H&C line from another line at the origin stations and interchange again to yet another line at the destination stations. 4) The non-interchange passengers: the passengers just travel on the H&C line without interchanging.

4.2.3. The influence of the station gates

Passengers enter or leave the metro system through station gates. One station may have several platform exits and station gates, and each exit may link exclusively to a station gate, like Aldgate East Station, where there are two platform exits and two station gates. Thus, there are two target cars for passengers travelling from Hammersmith Station to Aldgate East Station, because they may leave Aldgate East Station through different gates.

We assume that all the passengers using the destination-based choice for their boarding car can reach the target cars, and an ideal destination-based choice sub-model of the number of boarders on each car of a train is shown as follows:

$$Bd^i = B^i - Bo^i = (1 - \alpha)B^i \quad (6)$$

$$Bd_m^{i,j} = Bd^i \cdot (P_1^{i,j} \cdot G_{1,m}^{i,j} + P_2^{i,j} \cdot G_{2,m}^{i,j} + P_3^{i,j} \cdot G_{3,m}^{i,j} + P_4^{i,j} \cdot G_{4,m}^{i,j}) \quad (7)$$

$$Bd_m^i = \sum_{j=i+1}^{29} Bd_m^{i,j} \quad (8)$$

where

Bd^i is the number of boarders of a train at station i who use the destination-based choice for boarding cars;

Bd_m^i is the number of boarders of a train at station i , who board car m using the destination-based choice;

$Bd_m^{i,j}$ is the number of boarders of a train at station i , who board car m using the destination-based choice and travel to station j ;

$P_1^{i,j}, P_2^{i,j}, P_3^{i,j}, P_4^{i,j}$ are the percentages of passengers travelling from station i to j , who belong to non-interchange, interchange-in, interchange-out and interchange-in-out passenger types respectively. $P_1^{i,j}, P_2^{i,j}, P_3^{i,j}, P_4^{i,j}$ can be calculated based on RODS;

$G_{1,m}^{i,j}, G_{2,m}^{i,j}, G_{3,m}^{i,j}, G_{4,m}^{i,j}$ are the allocation coefficients for car m if it is the target car of passengers travelling from station i to j , who belong to non-interchange, interchange-in, interchange-out and interchange-in-out passenger types respectively.

$G_{1,m}^{i,j}, G_{2,m}^{i,j}, G_{3,m}^{i,j}, G_{4,m}^{i,j}$ are equal to 1 if car m is the only target car; they are equal to 0.5 if car m is one of the two target cars; they are equal to 0 if car m is not the target car.

4.3. The modified destination-based choice sub-model of the number of boarders on each car of a train taking into consideration the waiting time before the arrival of the train

In reality, it is not possible that all the passengers reach their target cars on the platform because of the waiting time before the arrival of the train. In order to avoid missing a train, passengers have to board the cars closest to them when the train stops, even if they are not their target cars. Therefore, the modified destination-based choice sub-model of the number of boarders on each car of a train is shown as follows:

$$Bd_m^{i'} = \sum_{e=1}^{en^i} \left(Bd_m^{i'} \right)_e \quad (9)$$

$$\left(Bd_m^{i'} \right)_e = \sum_{j=i+1}^{29} \left(Bd_m^{i,j'} \right)_e \quad (10)$$

where

$Bd_m^{i'}$ is the number of boarders of a train at station i , who board car m using the destination-based choice in the modified destination-based choice sub-model;

$\left(Bd_m^{i'} \right)_e$ is the number of boarders of a train at station i , who reach the platform through entrance e and board car m using the destination-based choice in the modified destination-based choice sub-model;

$\left(Bd_m^{i,j'} \right)_e$ is the number of boarders of a train at station i , who reach the platform through entrance e , board car m using the destination-based choice and travel to station j in the modified destination-based choice sub-model.

$\left(m_0^i \right)_e$ denotes the car number of the car nearest to the platform entrance e at station i .

If $m = (m_0^i)_e$, then,

$$\left(Bd_m^{i,j'}\right)_e = \left(Bd_m^{i,j}\right)_e + \frac{|x_e - x_{m-1,3}|}{v \cdot h^i} \cdot \sum_{m_x=1}^{m-1} \left(Bd_{m_x}^{i,j}\right)_e + \frac{|x_e - x_{m+1,1}|}{v \cdot h^i} \cdot \sum_{m_x=m+1}^7 \left(Bd_{m_x}^{i,j}\right)_e \quad (11)$$

where

m_x is the car number of a car other than car m , $m_x \in [1,7]$;

v is the average walking velocity of passengers, which is 0.5 m/s in this model;

h^i is the headway at station i , which is 300 s;

$\left(Bd_m^{i,j}\right)_e$ is the number of boarders of a train at station i , who reach the platform through entrance e , board car m using the destination-based choice and travel to station j in the ideal destination-based choice sub-model;

$\left(Bd_{m_x}^{i,j}\right)_e$ is the number of boarders of a train at station i , who reach the platform through entrance e , board car m_x using the destination-based choice and travel to station j in the ideal destination-based choice sub-model;

$x_{m-1,3}$ is the coordinate of the middle line of door 3 of car $m-1$;

$x_{m+1,1}$ is the coordinate of the middle line of door 1 of car $m+1$.

If $m > (m_0^i)_e$, then,

$$\left(Bd_m^{i,j'}\right)_e = \left(Bd_m^{i,j}\right)_e - \frac{|x_e - x_{m,1}|}{v \cdot h^i} \cdot \left(Bd_m^{i,j}\right)_e + \frac{|x_{m,1} - x_{m+1,1}|}{v \cdot h^i} \cdot \sum_{m_x=m+1}^7 \left(Bd_{m_x}^{i,j}\right)_e \quad (12)$$

where

$x_{m,1}$ is the coordinate of the middle line of door 1 of car m ;

$x_{m+1,1}$ is the coordinate of the middle line of door 1 of car $m+1$.

If $m < (m_0^i)_e$, then,

$$\left(Bd_m^{i,j'}\right)_e = \left(Bd_m^{i,j}\right)_e - \frac{|x_e - x_{m,3}|}{v \cdot h^i} \cdot \left(Bd_m^{i,j}\right)_e + \frac{|x_{m,3} - x_{m-1,3}|}{v \cdot h^i} \cdot \sum_{m_x=1}^{m-1} \left(Bd_{m_x}^{i,j}\right)_e \quad (13)$$

where

$x_{m,3}$ is the coordinate of the middle line of door 3 of car m ;

$x_{m-1,3}$ is the coordinate of the middle line of door 3 of car $m-1$.

$$\left(Bd_m^{i,j}\right)_e = Bd^i \cdot \left[\left(P_1^{i,j}\right)_e \cdot G_{1,m}^{i,j} + \left(P_2^{i,j}\right)_e \cdot G_{2,m}^{i,j} + \left(P_3^{i,j}\right)_e \cdot G_{3,m}^{i,j} + \left(P_4^{i,j}\right)_e \cdot G_{4,m}^{i,j} \right] \quad (14)$$

where

$\left(P_1^{i,j}\right)_e, \left(P_2^{i,j}\right)_e, \left(P_3^{i,j}\right)_e, \left(P_4^{i,j}\right)_e$ are the percentages of passengers reaching the

platform through entrance e and travelling to station j from station i , who belong to non-interchange, interchange-in, interchange-out and interchange-in-out passenger types respectively. They can be calculated based on RODS.

4.4 The integrated model of the number of boarders onto each car of a train

Combining the origin-based choice sub-model and the modified destination-based choice sub-model, the integrated model can be developed as follows:

$$B_m^i = Bo_m^i + Bd_m^{i'} \quad (15)$$

4.5 The model of the number of alighters from each car of a train

Since the influence of passenger on-train movement can be neglected, passengers will alight from the same car onto which they boarded. That is, the number of alighters from a car at a certain station is based on the number of boarders on that car at the previous stations. Hence, the model of the number of alighters from each car of a train can be developed as:

$$A_m^i = Ao_m^i + Ad_m^i \quad (16)$$

where

A_m^i is the number of alighters from car m at station i ;

Ao_m^i is the number of alighters from car m at station i using the origin-based choice;

Ad_m^i is the number of alighters from car m at station i using the destination-based choice.

$$Ao_m^i = \sum_{k=1}^{i-1} Bo_m^k \cdot (P_1^{k,i} + P_2^{k,i} + P_3^{k,i} + P_4^{k,i}) \quad (17)$$

where

Bo_m^k is the number of boarders at station k , who board car m using the origin-based choice;

$P_1^{k,i}, P_2^{k,i}, P_3^{k,i}, P_4^{k,i}$ are the percentages of passengers travelling from station k to i , who belong to non-interchange, interchange-in, interchange-out and interchange-in-out passenger types respectively.

$$Ad_m^i = \sum_{k=1}^{i-1} Bd_m^{k,i'} \quad (18)$$

$$Bd_m^{k,i'} = \sum_{e=1}^{en^k} (Bd_m^{k,i'})_e \quad (19)$$

where

$Bd_m^{k,i'}$ is the number of boarders of a train at station k , who board car m using the destination-based choice and travel to station i in the modified destination-based choice sub-model;

$(Bd_m^{k,i'})_e$ is the number of boarders of a train at station k , who reach the platform through entrance e , board car m using the destination-based choice and travel to station i in the modified destination-based choice sub-model.

5. The model calibration and evaluation

5.1 Model calibration

Assuming the percentage of passengers using the origin-based choice for the boarding cars (α) and the total number of boarders of a train at station i (B^i) are known, the number of boarders on car m of a train at station i ($Bd_m^{i'}$) in the modified destination-based choice sub-model can be directly measured based on RODS and the layouts of the metro platforms on the H&C line.

For the origin-based choice sub-model, the parameters a_e^i and b_e^i should be calibrated. The calibration of these two parameters can be complicated at a station with more than one entrance on the platform. However, at a station with only one entrance on the platform, and where the entrance is at the end of the platform, such as at Hammersmith Station and Goldhawk Road Station, the calibration can be simplified.

At that type of station, the number of boarders who reach the platform through platform entrance e and board car m through door n using the origin-based choice, $(Bo_{m,n}^i)_e$, is equal to the number of boarders onto car m through door n , $Bo_{m,n}^i$.

Moreover, the coordinates of the platform entrance e , x_e , are equal to 0 m. Thus, the equation containing a_e^i and b_e^i can be simplified as follows:

$$Bo_{m,n}^i = a_e^i x_{m,n} + b_e^i \quad (20)$$

The coordinates of the middle line of door n of car m , $x_{m,n}$ are fixed for each train on the H&C line. In order to calibrate the parameters, the actual number of boarders onto car m through door n of a train at station i is needed, which can be provided by the loadweigh.

Based on the loadweigh, the difference between the number of boarders and alighters at a certain station can be measured. We can estimate the actual number of boarders on car m of a train at station i using the equation as follows:

$$\left(B_m^i\right)_{actual} = In_m^i + A_m^i \quad (21)$$

where

$\left(B_m^i\right)_{actual}$ is the actual number of boarders on car m at station i ;

In_m^i is the difference between the number of boarders and alighters on car m at station i .

It is assumed that the actual number of boarders on car m through door n of a train at station i , $\left(B_{m,n}^i\right)_{actual}$, is one third of $\left(B_m^i\right)_{actual}$ as shown below.

$$\left(B_{m,n}^i\right)_{actual} = \frac{\left(B_m^i\right)_{actual}}{3} \quad (22)$$

Then, using the least square method to do regression to fit the linear passenger distribution, a_e^i and b_e^i can be calculated as shown below.

$$a_e^i = \frac{N_{door} \cdot \sum_{m=1}^7 \sum_{n=1}^3 x_{m,n} \cdot \left(B_{m,n}^i\right)_{actual} - \sum_{m=1}^7 \sum_{n=1}^3 x_{m,n} \cdot \sum_{m=1}^7 \sum_{n=1}^3 \left(B_{m,n}^i\right)_{actual}}{N_{door} \cdot \sum_{m=1}^7 \sum_{n=1}^3 \left(x_{m,n}\right)^2 - \left(\sum_{m=1}^7 \sum_{n=1}^3 x_{m,n}\right)^2} \quad (23)$$

$$b_e^i = \frac{\sum_{m=1}^7 \sum_{n=1}^3 (x_{m,n})^2 \cdot \sum_{m=1}^7 \sum_{n=1}^3 (B_{m,n}^i)_{actual} - \sum_{m=1}^7 \sum_{n=1}^3 x_{m,n} \cdot \sum_{m=1}^7 \sum_{n=1}^3 x_{m,n} \cdot (B_{m,n}^i)_{actual}}{N_{door} \cdot \sum_{m=1}^7 \sum_{n=1}^3 (x_{m,n})^2 - \left(\sum_{m=1}^7 \sum_{n=1}^3 x_{m,n} \right)^2} \quad (24)$$

where

N_{door} is the total number of doors of a train on the H&C line, which is 21.

5.2 Evaluating the performance of the integrated model using the GEH statistic

The GEH statistic is an empirical formula that judges the fitness of a traffic assignment model (Ortuzar and Willumsen 2011). It overcomes the problems, which are caused by the variations in traffic flows in a network, that are encountered when using the percentage of the difference between observed and modelled values to evaluate the performance of a traffic assignment model. For example, a 10% difference in a large flow (like a motorway) may indicate the need to construct an extra lane, whereas a 10% difference in a low flow (local link) may not be that significant. However, the GEH statistic is less sensitive to this issue. A GEH index less than 5.0 is considered a good match between the modelled and observed value (Transport for London 2010).

Essentially, the model developed in this study is used to assign passengers to each car after inputting the total number of boarders of a train at a certain station. Therefore, the GEH statistic is used to compare the actual number of boarders on car m of a train at station i , $(B_m^i)_{actual}$, with the modelled number of boarders on car m at station i , B_m^i , shown as follows.

$$GEH = \sqrt{\frac{\left[(B_m^i)_{actual} - B_m^i \right]^2}{0.5 \cdot \left[(B_m^i)_{actual} + B_m^i \right]}} \quad (25)$$

5.3 Estimation of the percentages of passengers using the origin-based and the destination-based strategy for the boarding car

In sub-section 5.1, in order to calibrate the origin-based choice sub-model and calculate the modelled number of boarders on car m at station i (B_m^i), a value for the percentage of passengers using the origin-based choice for the boarding car (α) first has to be set. Therefore, B_m^i is a function of α , $B_m^i = f(\alpha)$. Then, the GEH statistic is used to compare the modelled and actual number of boarders on car m of a train at station i . Thus, the GEH value is also a function of α , $GEH = f(\alpha)$. The α value which corresponds to the lowest value of the GEH is considered as the estimated percentage of passengers using the origin-based choice. Consequently, the percentage of passengers using the destination-based choice ($1-\alpha$) can also be estimated.

6. Results and performance of the model

As shown in Figure 2, the cars with the largest number of boarders are the second-closest car to the platform entrances at both stations, i.e. Car 6 at Hammersmith Station and Car 2 at Goldhawk Road Station, respectively. It means the waiting area around the second-closest car to the entrances is the most crowded. At Hammersmith Station, the number of boarders at Car 2 is larger than those at Car 3 and 4, although the distances from the platform entrance to Car 3/4 are shorter than that to Car 2.

The performance of the model on the eastbound platform at Hammersmith Station and Goldhawk Road Station was tested. If we assume all the passengers use the destination-based choice for boarding cars, the passenger distribution on the platform corresponding to the locations of train cars fits the trend of the actual passenger distribution at both stations, as shown in Figure 2. Therefore, the destination-based choice for boarding cars has a great influence on the passenger distribution on the metro platform.

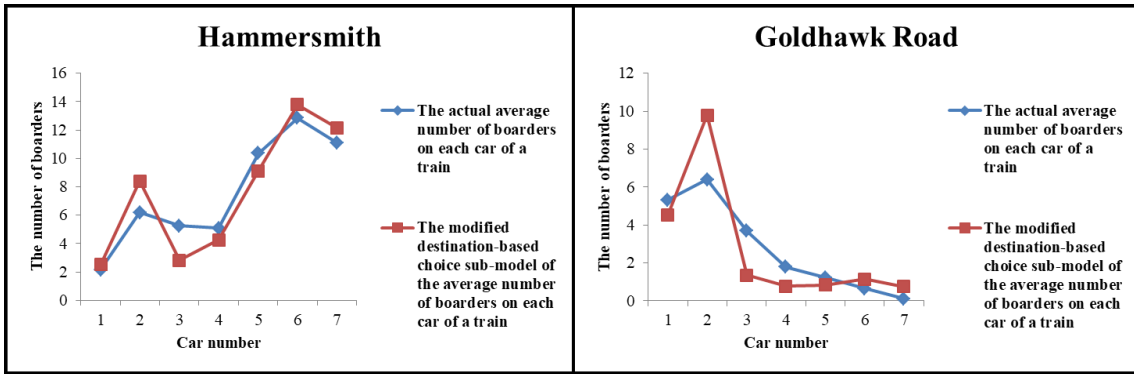


Figure 2: Comparison of the actual average number of boarders on each car and the estimated number in the destination-based choice sub-model

Combining the origin-based choice sub-model with the destination-based choice sub-model, the integrated model results of the average number of boarders on each car of a train at both stations can be calculated, as shown in Figure 3 and Table 1.

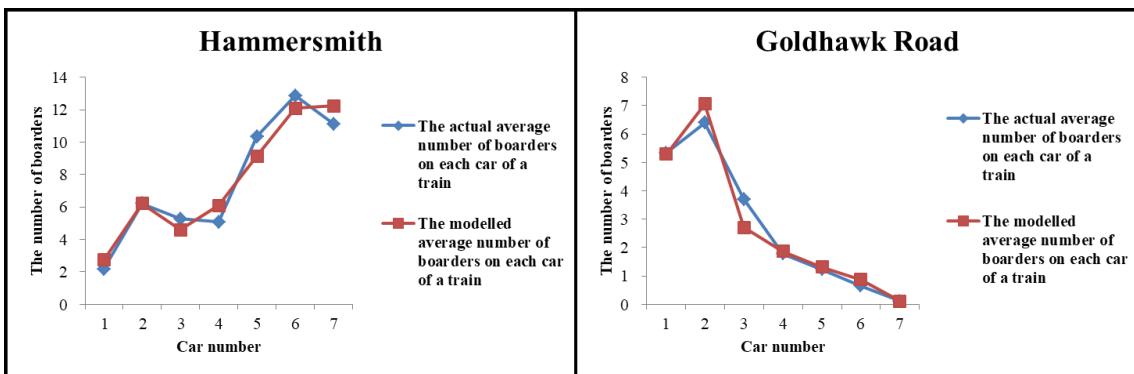


Figure 3: Comparison of the actual and modelled average number of boarders on each car

Table 1: The results of the model performance

	Car 1	Car 2	Car 3	Car 4	Car 5	Car 6	Car 7
The GEH at Hammersmith	0.35	0	0.30	0.43	0.39	0.22	0.32
The GEH at Goldhawk Road	0.02	0.25	0.55	0.06	0.07	0.24	0

The largest difference between the actual and modelled average number of boarders on a car of a train is less than 1.3 passengers. The GEH index is less than 1, which suggests the model performance is good and that the two boarding strategies may well explain the mechanism of the passenger distribution on the platform at the stations when there are an adequate number of seats on the train.

The highest possible percentages of passengers using the destination-based choice are calculated as 44.5% and 44.6% at Hammersmith Station and Goldhawk Road Station respectively. It is suggested that nearly half of the passengers in the morning peak times choose their boarding cars by taking into consideration the layouts of the stations where they would alight from the train.

7. Discussion

The model developed in this study, which considers the two boarding strategies (i.e. the origin-based and destination-based choice), showed good performance in predicting the passenger distribution on the platform at the two selected stations. As Figure 2 shows, there was a trend at both studied stations that the congestion level of the platform decreased relative to the increase in the distance from the platform entrance. This was the same result as found in the observations of previous research (Szplett and Wirasinghe 1984; Lam, Cheung, and Lam 1999; Wu, Yang, and Yuan 2010; Krstanoski 2014). In addition, at the two selected stations, it was the second-closest car to the platform entrance (Car 6 at Hammersmith Station and Car 2 at Goldhawk Road Station) that saw the largest number of boarding passengers, which was similar to the findings of Lam, Cheung, and Lam (1999). At Hammersmith Station, there was another passenger cluster around Car 2 between Car 1 and Car 4, because this car was the closest car to the platform exits at the destination stations of 31.7% of the passengers who boarded at Hammersmith. In contrast, the share of Hammersmith boarding passengers travelling to stations whose exits are close to Cars 3 and 4 were 0.4% and 4.6%, respectively, and this may be a reason why Cars 3 and 4 did not have many boarding passengers at Hammersmith Station, although the distances from the platform entrance to Car 3/4 are shorter than that to Car 2. Szplett and Wirasinghe (1984) also observed two clusters of passengers waiting on a platform with one entry. It is speculated that these clusters are,

as is seen at Hammersmith Station in this study, a result of the platform layouts of the destination stations of the passengers of the observed station and the passengers' strategy to minimise the walking distance at their destination stations.

It should be noted that, with regard to the variables that may affect passenger choice for boarding cars, this study considered only the layouts of the platforms at the passengers' boarding and alighting stations. There may be many other variables. For example, crowding on the platform and/or on the train can be another variable. In fact, the two stations selected for the case study in this research are the first two stations on the H&C line, where passengers usually can find seats. Passenger behaviours of choosing boarding cars were rarely affected by platform or/and train crowding. This may be a reason why the percentages of passengers using the destination-based choice are almost the same at both of these stations.

In addition to crowding, Wu, Yang, and Yuan (2010) found that obstructions to passengers' vision can affect passenger choice of boarding cars, because such obstructions can hinder passengers from finding less crowded areas on the platform to stand. Weather may also influence passengers' boarding at outdoor stations, like Latimer Road on the H&C line, where the roof cannot cover the whole platform. Passengers may prefer to stand in the area under the roof when it is raining. Furthermore, passengers' boarding strategies can depend on their mobility (related to age, health conditions, etc.), time of the day (passengers may use different boarding strategies in the morning peak, off-peak and evening peak hours) and the surrounding environment (an unpleasant smell, a person behaving oddly, etc.). In fact, these variables can be categorised into three groups: platform-related, train-related, and passenger-related, and each group may include many variables. Although this study took the decision to use two variables only (platform layouts and boarding and alighting

stations), the limitation that this study does not consider other factors needs to be acknowledged and further research is required to address this.

8. Conclusions

Uneven passenger distribution on the platform can lead to a longer train dwell time and accordingly impact the network capacity. It is important to study the underlying mechanism causing uneven passenger distribution before taking action to reduce it.

Because of the fact of passengers choosing boarding cars using a destination-based choice, the design of a platform layout should not only consider the platform capacity to meet the demand of passenger volumes, but also needs to refer to the passengers' OD matrix (which is calculated from the results of demand forecasting at the planning stage) to coordinate the locations of the platform entrances and exits at each station of a metro line, so that uneven passenger distribution on the platform can be reduced. For example, if of all the passengers starting journeys at station A the proportion of passengers travelling to station B is the largest, the platform entrance at station A and the platform exit at station B should not be close to the same car. Meanwhile, sufficient walking space on the platform should be incorporated into the design to enable passengers to walk along it comfortably. In addition, designing more than one entrance to the platform can help make passenger distribution tend to even out. The models can be used by metro operation planners to optimise the design of metro platform layouts.

In terms of station operation, metro staff can inform the passengers of the least crowded car(s) that are away from the platform entrances when the train stops. This may motivate some passengers to walk along the platform, instead of clustering near the entrances.

Acknowledgements

The authors thank the support of Transport for London. The results of this study does not reflect the official opinions of Transport for London. Jiping Fang acknowledge the financial support of the China Scholarship Council (Grant No. 201507090067).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Jiping Fang acknowledge the financial support of the China Scholarship Council [grant number 201507090067].

References

- BMBS (Beijing Municipal Bureau of Statistics). 2018. "Statistical Communiqué of Beijing Municipality on the 2017 Economic and Social Development." February 25. http://www.bjstats.gov.cn/zxfb/201802/t20180225_393332.html.
- Bian, Bomin, Ning Zhu, Shuai Ling, and Shoufeng Ma. 2015. "Bus service time estimation model for a curbside bus stop." *Transportation Research Part C: Emerging Technologies* 57:103-21.
- Fujiyama, T, J. Nowers, and N. Tyler. 2008. "Investigation into train dwell time, research report submitted to UK Department for Transport."
- Kim, H., S. Kwon, S.K. Wu, and K. Sohn. 2014. "Why Do Passengers Choose a Specific Car of A Metro Train During the Morning Peak Hours?" *Transportation Research Part A: Policy and Practice* 61 (2014): 249-258.
- Krstanoski, N. 2014. "Modelling Passenger Distribution on Metro Station Platform." *International Journal for Traffic & Transport Engineering* 4(4): 456-465.

- Lam, W.H., C.Y. Cheung, and C.F. Lam. 1999. "A Study of Crowding Effects at the Hong Kong Light Rail Transit Stations." *Transportation Research Part A: Policy and Practice* 33(5): 401-415.
- Lu, J., and F. Dong. 2010. "Statistical Study of Passenger Distribution on Urban Rail Transit Platform Before the Train Stop." *Urban Mass Transit* 13(7): 53-56.
- NHS Digital. 2016. "Health Survey for England 2015 Trend Tables." December 14. <http://digital.nhs.uk/catalogue/PUB22616>
- Ortuzar, J. de D., and L.G. Willumsen. 2011. *Modelling Transport*. New York City: John Wiley & Sons.
- Szplett, D., and S.C. Wirasinghe. 1984. "An Investigation of Passenger Interchange and Train Standing Time at LRT Stations:(i) Alighting, Boarding and Platform Distribution of Passengers." *Journal of Advanced Transportation* 18(1), pp.1-12.
- Transport for London. 2018. "Annual Report and Statement of Accounts 2017/18 draft." <http://content.tfl.gov.uk/tfl-draft-annual-report-2017-2018.pdf>.
- Transport for London. 2010. "Traffic Modelling Guideline." <http://content.tfl.gov.uk/traffic-modelling-guidelines.pdf>
- Wirasinghe, SC, and David Szplett. 1984. "An investigation of passenger interchange and train standing time at LRT stations:(ii) estimation of standing time." *Journal of advanced transportation* 18 (1):13-24.
- Wu, F., Y. Yang, and Z. Yuan. 2010. "Waiting Location Choice of Passengers in Urban Rail Transit Platform During the Train Stop." *Urban Mass Transit* 13(9): 52-56.
- Wu, Y.Y., J. Rong, X.M. Liu, and Z.H. Wei. 2012. "Passengers Distribution in Urban Rail Transit Platform Before Vehicle Arrival." *Journal of Beijing University of Technology*, 38(6): 875-879.