Zhang, Q, Liu, X, Spurgeon, S and Yu, DL

A two-layer modelling framework for predicting passenger flow on trains: A case study of London underground trains

Article

Zhang, Q, Liu, X, Spurgeon, S and Yu, DL (2021) A two-layer modelling framework for predicting passenger flow on trains: A case study of London underground trains. Transportation Research Part A: Policy and Practice, 151. pp. 119-139. ISSN 0965-8564

A Two-layer Modelling Framework for Predicting Passenger Flow on Trains: A Case Study of London Underground Trains

Qian Zhang[1*], Xiaoxiao Liu[2], Sarah Spurgeon[3] and Dingli Yu[1]

[1]*Department of Electronics and Electrical Engineering, Liverpool John Moores University, Liverpool, UK.*

[2]*Department of Maritime and Mechanical Engineering, Liverpool John Moores University, Liverpool, UK.*

[3]*Department of Electronic and Electrical Engineering, University College London, London, UK.*

[*]Corresponding author. E-mail: Q.Zhang@ljmu.ac.uk; Department of Electronics and Electrical Engineering, Liverpool John Moores University, 3 Byrom Street, Liverpool, L3 3AF, UK.

Abstract: A model that anticipates the passenger flow on trains will help passengers to avoid overcrowded trains in their journey planning. Such a model will also help rail industry to understand the current use of train capacity and plan the distribution of rolling stock, personnel and facilities. However, the existing studies only developed the models for forecasting the passenger flow in stations, which cannot reflect the true passenger number on trains. In this paper, a hierarchical modelling framework for passenger flow prediction is proposed. It includes two layers of fuzzy models, where a global model is used to predict for ordinary circumstances and a number of local models are used to predict the variations in passenger number due to specific factors, such as events and weather. A new data sifting method is proposed to obtain the most informative and representative data for model training, which greatly improves the modelling efficiency. The proposed method is then validated using a case study of forecasting the passenger flow of London Underground trains.

Keywords: passenger flow; train; fuzzy modelling; data sifting; event; weather

1. Introduction

Improving passengers' experience is a crucial challenge for the rail transport. One of the key factors that affect the passenger experience is overcrowding, and the crowding level on trains may even affect the path/train choice of passengers (Pel et al. 2014, Kim et al. 2015). If an accurate passenger flow model is available to predict the crowding level on trains, it can help rail passengers to strategically plan their journeys to avoid overcrowded trains, by choosing alternative means of transportation or changing the travel time. It can also help rail operators and administrators to appropriately allocate train capacity, personnel and facilities and adjust the service timetable.

In the past ten years, there have been many models developed to predict the passenger flow in stations (Wei and Chen 2012; Jiang et al. 2014; Ding et al. 2018). More details about these studies can be found in the literature review of Section 2. However, to the best of our knowledge, there is no research that studied the passenger flow on trains and developed the relevant forecasting models. On the other hand, most of the existing models could only predict the passenger flow under regular conditions, and there are limited works investigated the variation of the passenger flow when large events happen (Pereira et al. 2015; Chen et al. 2020) or under different weather conditions (Wang et al. 2018).

In this research, we studied the historical data of a London Underground train line and proposed a data-driven modelling framework for predicting the passenger flow on trains. First, the elements that affect the passenger number were identified, including temporal factors (time in a day and day of the week), location, direction, local events and weather. We then proposed a hierarchical modelling framework including two layers of models. A global model was developed to predict the passenger flow under ordinary circumstances and a number of local models were developed to predict the changes in passenger number caused by specific factors, such as events and weather conditions. These models implement fuzzy rule-based systems (FRBSs) (Zadeh 1973) with a heuristic optimisation technique (Zhang and Mahfouf 2010) and a novel data sifting strategy. The heuristic optimisation technique provides FRBSs the ability of learning from data and the data sifting strategy improves the efficiency of model generation. Different with the time-series models that are suitable to predict the passenger flow for near future, such as the autoregressive (AR) or autoregressive integrated moving average (ARIMA) models, the FRBS is able to predict for any time point and does not rely on the observation of the current passenger flow. Different with other commonly-used non-parametric models, such as artificial neural networks (ANNs), the FRBS is more transparent due to its use of linguistic IF-THEN rules. Finally, the framework was successfully applied into the prediction of the passenger load on the London Underground trains.

This research would help the rail industry understand the factors affecting the passenger number on trains and be able to predict the impact of these factors. The developed predictive models can be integrated with existing journey planning tools to help travellers better plan their journeys to avoid overcrowded trains, where they may choose alternative means of transportation (private cars, taxis, buses, shared bicycles, walking, etc.) or slightly change the travel time. Combined with multi-criteria optimisation and decision making techniques, the developed models can be further exploited to consider the personalised preferences of passengers and automatically suggest the best decisions for journey planning. This research also has a potential to inform distribution of rolling stock, personnel and facilities in that a good predictor of passenger load could enable better optimisation of train capacity and other resources.

The novelty of this work lies in the following aspects: First, this work is a very early and timely research in modelling the passenger flow on trains. It considers not only the ordinary conditions but also the effect of events and weather. Second, this work employs FRBSs to map the relationship between model inputs and output, where FRBSs are more transparent than commonly-used black-box methods due to the presence of linguistic rules. Last, this work proposes a clustering-based data sifting method for selection of the most representative data for model training. This helps greatly reduce the time and resources used in model generation and is very valuable in the current data-rich era.

The remaining sections are organised in the following way. Section 2 reviews the existing methods of the passenger flow prediction, especially focusing on the applications in the rail transport. In Section 3, the data from London Underground trains are analysed and the key attributes that affect the passenger number are identified. Section 4 proposes a data-driven modelling method with a two-layer structure. A salient data sifting mechanism is invented to enhance the modelling efficiency. In Section 5, the proposed modelling approach is implemented to construct passenger flow models for London Underground services. Finally, conclusions and implications of this study are presented in Section 6.

2. Related Work

In the past two decades, there have been extensively studies carried out on passenger demand and traffic flow (Li and Sheng 2016; Melo et al. 2019), especially focusing on short-term traffic/passenger flow prediction. These studies range across road transport (Smith et al. 2002; Vlahogianni et al. 2014; Sheng and Sharp 2019), rail transport (Tsai et al. 2009; Li et al. 2017a), waterborne transport (Kim and Lee 2018; He et al. 2019) and air transport (Faraway and Chatfield 1998; Bao et al. 2012). There have been several review articles discussing the technologies and applications in the field of short-term traffic flow prediction for road transport (Vlahogianni et al. 2004, 2014; Karlaftis and Vlahogianni 2011).

Considering the modelling architecture, the relevant research can be grouped into two main categories: parametric modelling and non-parametric modelling. The parametric models include the ARIMA models (Lee and Fambro 1999; Williams and Hoel 2003; Tan et al. 2009; Kumar and Vanajakshi 2015), state space models (Stathopoulos and Karlaftis 2003), grey models (Hsu and Wen 1998), etc. The ARIMA model is an effective regressive time-series model and is the most widely used parametric method in traffic/passenger flow prediction. However, it assumes linear relationships among the time variables and thus suffers for dealing with the problems involving nonlinear relationships.

The non-parametric modelling approaches include ANNs (McFadden et al. 2001; Ishak et al. 2003; Chan et al. 2012), support vector machine (SVM) (Wu et al. 2004; Chen et al. 2012; Jiang et al. 2014), Kalman filter (Wang et al. 2007; Lippi et al. 2013), non-parametric regression (Clark 2003; Sun et al. 2014), wavelet analysis (He and Ma 2002, Jiang et al. 2005; Sun et al. 2015), etc. In these non-parametric approaches, ANNs have received the greatest interests and have been extensively implemented, due to their advantages of mapping nonlinear relationships and high accuracy. Various types of ANNs have been utilised, such as multi-layer perceptron networks (Zhang 2000; Smith et al 2002;), recurrent neural networks (van Lint et al. 2002; Ishak et al. 2003; Zhang et al. 2018a), radial basis function networks (Park et al 1998; Li et al 2017), spectral basis networks (Park et al. 1999), etc.

Besides the methods mentioned above, some research (Zeng et al. 2008; Li et al. 2014) tried to integrate both the parametric and non-parametric methods to achieve better performance. Several other studies (Chen and Wu 2012; Wei and Chen 2012; Jiang et al. 2014) combined the empirical mode decomposition (EMD) with the parametric and non-parametric approaches to further improve the modelling accuracy.

Compared with the traffic/passenger flow studies in road transport, the relevant research in rail transport is still in an emerging stage. Table 2-1 summarises some representative studies about the passenger flow prediction in rail transport, where most of these work arose since 2014. Several studies considered national railway services and predicted daily (Tsai et al. 2009; Jiang et al. 2014) or monthly (Milenkovic et al. 2018) passenger demand. Most of the research studied Metro systems due to the relatively easy access to the passenger flow data, which were generally collected from automated fare collection (AFC) systems. The existing models can only predict the passenger flow in stations or on platforms, and cannot accurately predict the passenger number on a certain train, especially in the case multiple train services share the same station or platform. However, the crowding level on a train is normally higher than that in a station and it is the more significant factor that affects the passenger experience. Though some approaches (Hörcher et al. 2017; Hänseler et al. 2020) were proposed to achieve the assignment of passengers to individual trains, there is still a lack of dedicated models that can directly predict the passenger flow on trains. In the research of this paper, we utilised the data collected from train loading sensors to anticipate the detailed passenger number on trains.

Table 2-1. An overview of the studies on rail passenger flow prediction.

Literature	Area	Case Study	Method	Data Source	Problem		
					Variable	Object	Condition
Tsai et al. 2009	National railway	Taipei-Kaohsiung, Taiwan	ANN	Daily tickets	Passenger demand	Railway	Ordinary
Wei and Chen 2012	Metro	Taipei, Taiwan	ANN and EMD	AFC	Passenger flow	Station	Ordinary
Leng et al. 2013	Metro	Beijing, China	Probability tree	AFC	Passenger flow	Station	Ordinary
Dou et al. 2014	National railway	Beijing-Jinan, China	FKNN	Tickets	Passenger flow	Railway	Ordinary
Sun et al. 2014	Metro	Beijing, China	NR	AFC	Passenger flow	Station	Ordinary
Jiang et al. 2014	National railway	Wuhan-Guangzhou, China	SVM and EMD	Daily tickets	Passenger demand	Station	Ordinary
Zhang et al. 2014	Metro	New York, USA	Network Kriging	AFC	Passenger demand	Station	Ordinary
Pereira et al. 2015	Metro and bus	Singapore	ANN	Internet and AFC	Passenger flow	Station	CE
Sun et al. 2015	Metro	Beijing, China	SVM and Wavelet	AFC	Passenger flow	Station	Ordinary
Ding et al. 2016	Metro	Beijing, China	GBDT	AFC	Passenger flow	Station	Ordinary
Jiao et al. 2016	Metro	Beijing, China	Kalman filter	AFC	Passenger flow	Station	Ordinary
Li et al. 2017a	Metro	Beijing, China	ANN	AFC	Passenger flow	Station	Ordinary, CE
Ni et al. 2017	Metro	New York, USA	ARIMA and LR	Social media and TU	Passenger flow	Station	CE
Ding et al. 2018	Metro	Beijing, China	ARIMA and GARCH	AFC	Passenger flow	Station	Ordinary
Gong et al. 2018	Metro	Sydney, Australia	ONMF	AFC	Passenger flow	Network	Ordinary
Li et al. 2018	Metro	Xi'an, China	ARIMA and SR	AFC	Passenger flow	Station	Ordinary
Ling et al. 2018	Metro	Shenzhen, China	ANN, SVM and GBDT	AFC	Passenger flow	Station	Ordinary
Milenkovic et al. 2018	National railway	Serbia	ARIMA	National Statistical Office	Passenger demand	Railway	Ordinary
Wang et al. 2018	Metro	Shanghai, China	LR	AFC	Passenger flow	Station	Ordinary, CW
Chen et al. 2019	Metro	Chengdu, China	ANN and EMD	AFC	Passenger flow	Station	Ordinary
Guo et al. 2019	Metro	Guangzhou, China	ANN and SVM	AFC	Passenger flow	Station	Ordinary
Jia et al. 2019	Metro	Guangzhou, China	ANN	AFC	Passenger flow	Station	Ordinary
Li et al. 2019a	Metro	Beijing, China	ANN	AFC	Passenger flow	Station	Ordinary
Liu et al. 2019a	Metro	Nanjing, China	ANN	AFC	Passenger flow	Station	Ordinary
Liu et al. 2019b	Metro	Beijing, China	Wavelet	AFC	Passenger flow	Network	Ordinary
Tang et al. 2019	Metro	Chongqing, China	ANN	AFC	Passenger flow	Station	Ordinary
Xiong et al. 2019	Metro	Beijing, China	ANN	AFC	Passenger flow	Station	Ordinary
Yang et al. 2019	Metro	Chongqing, China	ANN	AFC	Passenger flow	Station	Ordinary
Zhang et al. 2019a	Metro	Beijing, China	ANN	AFC	Passenger flow	Station	Ordinary
Zhao and Mi, 2019	National railway	Beijing-Shenzhen, China	ANN, SVM and Wavelet	Daily tickets	Passenger demand	Station	Ordinary
Zhu et al. 2019	Metro	Qingdao, China	ANN and SVM	AFC	Passenger flow	Station	Ordinary
Chen et al. 2020	Metro	Nanjing, China	ARIMA and GARCH	AFC	Passenger flow	Station	Ordinary, CE
Jing and Yin, 2020	National railway	Beijing, China	ANN	Daily tickets	Passenger demand	Station	Ordinary
Lu et al. 2020	Metro	Xi'an, China	ANN	AFC	Passenger flow	Station	Ordinary
Zhang et al. 2020	Metro	Xiamen, China	ANN	AFC	Passenger flow	Station	Ordinary
Zhao et al. 2020	Metro	Shanghai, China	ANN and STL	AFC	Passenger flow	Station	Ordinary
This research	Metro	London, UK	FRBS	Train loading	Passenger flow	Train	Ordinary, CE, CW

Method: ANN (artificial neural network), EMD (empirical mode decomposition), FKNN (fuzzy k-nearest neighbour), NR (non-parametric regression), GBDT (gradient boosting decision tree), LR (linear regression), GARCH (generalized autoregressive conditional heteroskedasticity), ONMF (online non-negative matrix factorization), SR (symbolic regression); STL (seasonal and trend decomposition using loess) **Data Source:** AFC (automated fare collection), TU (turnstile usage); **Condition:** CE (considering events), CW (considering weather)

Most of the research in the table considered the passenger flow under ordinary conditions and only several papers (Pereira et al. 2015; Li et al. 2017a; Ni et al. 2017; Chen et al. 2020) studied the effect of large events on the passenger flow. Pereira et al. (2015) and Ni et al. (2017) used Internet and social media data to predict the passenger arrivals using public transport, including Metro and buses. Li et al. (2017) and Chen et al. (2020) developed models for forecasting the non-regular passenger demands of Metro under special events. Wang et al. (2018) considered weather as one of the factors that affects the passenger flow of the Metro stations.

Among these research, ANN and ARIMA are the most frequently used methods, and other employed methods include EMD, SVM, Kalman filter, linear regression, non-parametric regression, gradient boosting decision tree (GBDT), generalized autoregressive conditional heteroscedasticity (GARCH), etc. The work in this paper is an early attempt to employ fuzzy rule-based systems as models. FRBSs are more flexible to use than the commonly used time-series models, as they do not rely on the observation of the current passenger flow and can predict for any time in future. FRBSs are more interpretable than the commonly used black-box methods due to the presence of linguistic rules. Several pieces of research (Zhang and Ye 2008; Li et al. 2016; Yu et al. 2019) have applied fuzzy sets and fuzzy systems into the traffic/passenger flow prediction in road transport. In the work of Dou et al. (2014), the fuzzy set concept was employed to improve the k-nearest neighbour (KNN) model in prediction of the passenger flow for a high-speed railway and was shown to outperform the original KNN model and the ARIMA model. However, it did not employ a FRBS as the main modelling structure and thus did not take the full advantage of a fuzzy system in modelling accuracy and transparency.

The difference between our research and other existing research lies in three aspects: First, our research models the passenger flow on trains; while other research focused on the passenger flow in stations or on platforms. Second, our research considers not only the ordinary conditions but also the effect of events and weather; while the majority of other research only considered ordinary conditions and only several studied the effect of either events or weather. Third, our research employs FRBSs in modelling, which can achieve a good balance between accuracy and transparency; while other research mainly employed either accurate ANN models or transparent ARIMA models.

3. Data Collection and Analysis

The train loading data studied in this research were collected by Transport for London (TfL), UK, from a London Underground train line, the Victoria line. The Victoria line is one of the busiest Underground lines running across central London from the south to the northeast. Figure 3-1 shows the map of the Victoria line and Table A-1 in appendix illustrates the abbreviations of its station names. The Victoria line carries 200 million passengers each year and thus generates a huge amount of data.

The collected data cover a period of one year that starts from 01 October 2014 and ends by 30 September 2015. The data of the train loading were measured by an automatic weighting system fitted on trains. They include all the loading information of each individual train at every station across the whole year. Some other relevant information, such as the train number, lead car number, train destination, direction and actual departure time at each station, was also provided. There are more than 4.86 million records in total.

Figure 3-1: Map of the Victoria line.

The loading of each train is returned as a percentage of total crush loading capacity in the range between 0 and 1. This index value is called passenger load status. It is an approximation based on the averaged measurements of the secondary suspension pressure transducer across 8 cars on each train. 0% loading represents a train with no passengers on it and 100% loading represents a crush laden train which has 1431 passengers on it, where it assumes an average weight per passenger is 72kg. At the current stage, we do not consider the uncertainty of the assumption and conversion. The following equation is used to calculate the passenger number from a given train load status:

$$Passenger\ number(t) = \frac{(Passenger\ load\ status(t)) \times 1431}{100}. \quad (1)$$

For analysis and modelling, the passenger load status is averaged in each fixed interval of 30 minutes, between 5:00 am and 1:00 am next day (daily continuous period for services running). The interval of 30 minutes is an appropriate setting for the case study, which ensures that there are at least several trains running during each interval. Public holidays, such as Christmas and New Year days, are excluded from the analyses of normal workdays.

Weather data and events data were also provided by TfL. The weather data during the aforementioned study period were recorded in every 15 minutes for 24 hours from nineteen weather stations in the London area. They contain temperature, rain state, rain intensity, snow state, etc. and there are more than 650 thousands weather records in the data set. More than three thousands events happened in London and the information about their locations, start time and finish time were collected and included in a data base.

The following paragraphs assess the attributes that make important contribution to the number of passengers. The potential attributes include temporal variables (month in a year, day of the week and time in a day) (Li et al. 2017a; Ding et al. 2018; Liu et al. 2019b), positional variables (station and travel direction) (Ding et al. 2018; Li et al. 2018; Liu et al. 2019a), weather (Wang et al. 2018) and events (Li et al. 2017a; Ni et al. 2017; Chen et al. 2020). In the following analyses, the passenger load in a certain day will be compared with the annual average of the passenger load. In calculation of the annual average, the abnormal days, such as bank holidays and the days when extreme weather, big events or long service disruption happens, are excluded. In this study, 122 days out of 365 days are identified as abnormal and have been excluded from the calculation of the annual average. In the following analyses, the passenger load at a certain station refers to the passenger load when trains approach that station.

As demonstrated in Figure 3-2, the passenger flow varies a lot in a single day. For weekdays, its pattern is highly similar, where two peaks can be observed for most stations. The first peak generally occurs in the morning between 7:30 and 9:30 and the second peak is in the afternoon between 17:00 and 19:00. Normally, the passenger flow in the morning peak and/or the afternoon peak could be 2-3 times of that in off-peak hours. The weekend shows different trends compared to weekdays and there is no distinct morning and afternoon peaks. High passenger load often appears between 11:00 and 19:00. The passenger load on Saturday is generally higher than that on Sunday. It is also noticed that there is an increase in passenger flow during the late nights on Friday and Saturday due to entertainment and social activities. These observation shows that day of the week and time in a day have a significant impact on the passenger flow, as indicated in (Liu et al. 2019b). We also investigated the variation of the passenger flow in different seasons. It was found that the mean passenger flow keeps a very similar trend across four seasons.

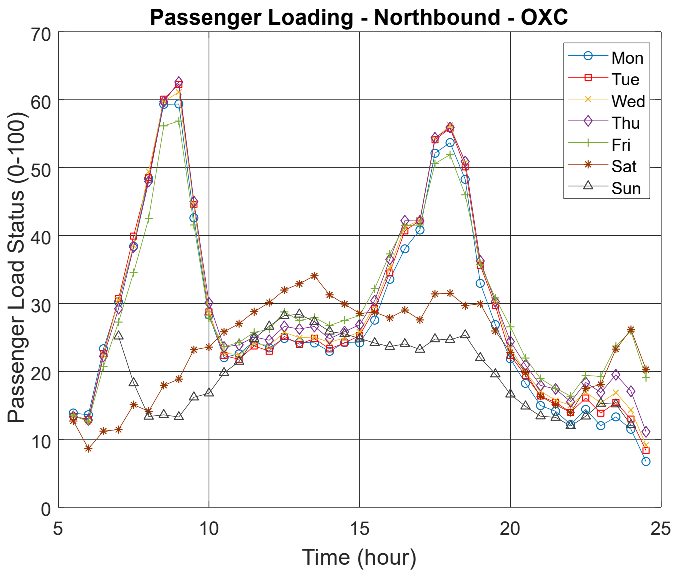
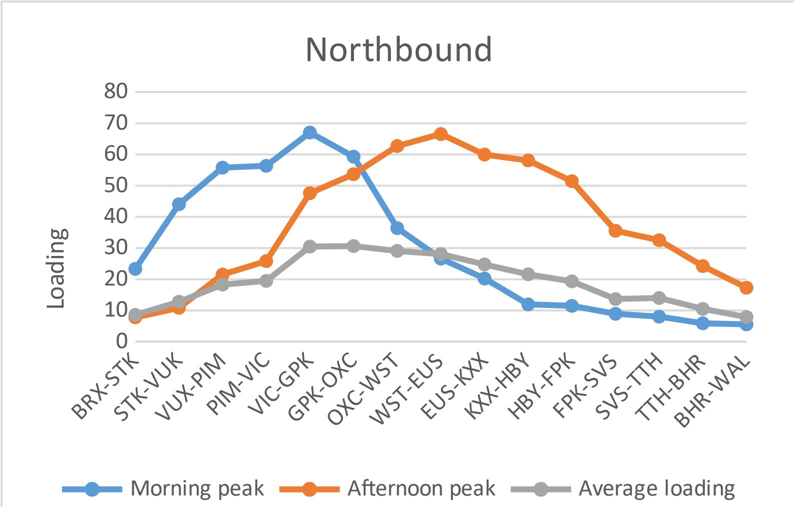
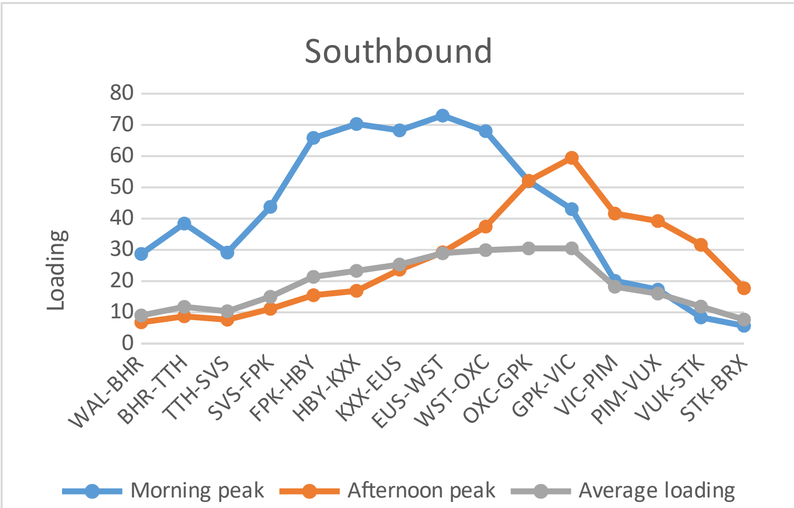


Figure 3-2: Mean passenger flow for different days of a week at Oxford Circus.

From a starting station to an end station, the mean passenger load shows a trend of first increasing and then decreasing along the Victoria line. Figure 3-3 shows how the passenger load level (0-100) changes with respect to different stations for both northbound and southbound trains. It can be observed that the northbound trains and the southbound trains at the same station produce very different figures. We can also find that the maximum passenger load normally appears in the morning peak of the southbound trains between Highbury & Islington and Warren Street. From the average curve, one can see that the passenger flow at the stations in the city centre could be 2-3 times of that at the stations close to the termini. The different scenarios observed from different stations and different running directions show that positional variables largely affect the passenger number.



(a)



(b)

Figure 3-3: Variation of mean passenger load for morning peak, afternoon peak and day average at different stations: (a) northbound trains and (b) southbound trains.

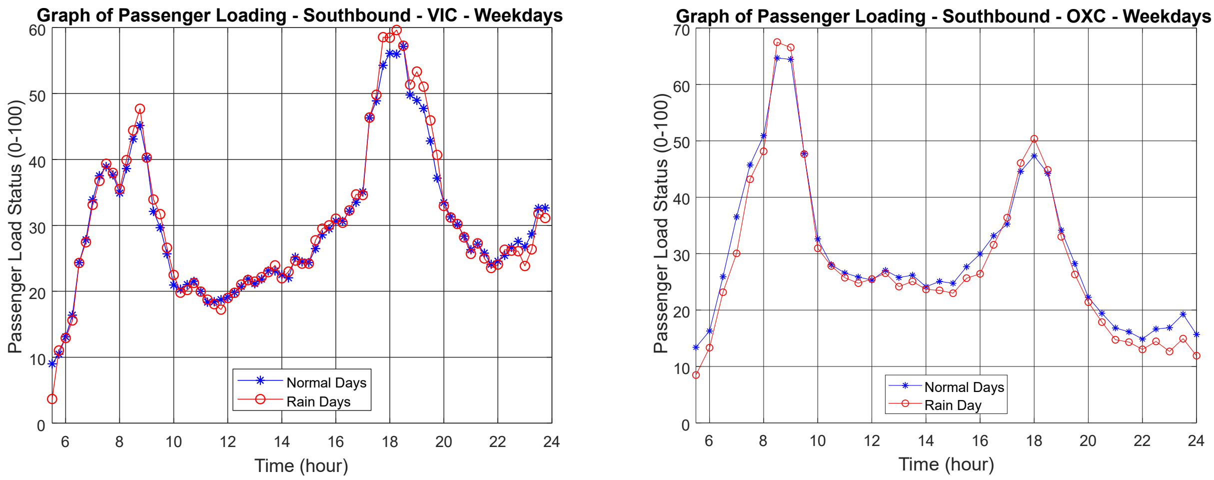
Two statistical measures, Pearson's correlation coefficient (for linear relationship detection) and correlation ratio (for non-linear relationship detection), were employed to study the significance of different variables to the passenger load. From Table 3-1, it can be seen that time in a day and station ID have relatively large correlation with the passenger load and their relationships are non-linear, which is consistent with the observations from the previous figures.

Table 3-1: Correlation between passenger load and other relevant factors.

	Time in a day	**Day of a week**	**Month**	**Station ID**
Corr coef *r*	-0.080*	-0.105*	-0.043*	-0.060*
Corr ratio	0.352	0.154	0.050	0.490

*significance level of 1%.

Weather condition is considered to closely associate with human lives. It could make an impact on decisions and activities of human, including choice of transport and trip planning. Some adverse weather conditions, such as heavy rain and snow, may be important factors that affect the passenger number of Underground trains. Figure 3-4 shows two examples of the passenger flow on the raining days. It can be seen that the passenger load at peak time is slightly higher than the yearly average value and the passenger load at off-peak time is slightly lower than the yearly average value. In the morning and afternoon peak time, the majority of the travellers would be commuters, who may choose to change their travel modes from walking, cycling and buses to subway, which is less weather affected. Wu and Liao (2020) found a similar scenario that the travellers tended to choose subway or cars rather than cycling and buses under inclement weather conditions. In the off-peak time, there are many people travelling for leisure purposes. They have the flexibility and may consider cancelling or delaying their inessential trips, where some research found that the leisure trips are more sensitive to adverse weather conditions (Arana et al. 2014). However, the effect of weather on subway passenger flow is relatively small compared to other means of transport due to the reliability of rail transport (Nosal and Miranda-Moreno 2014; Wu and Liao 2020). Some research (Liu et al. 2019a) even found heavy rain had a negligible impact on the passenger flow of the whole subway network.



(a) (b)

Figure 3-4: Effect of rain on passenger flow at Victoria: (a) 04 Nov 2014, Tuesday, maximum intensity 9.1 mm/h and (b) 13 Aug 2015, Thursday, maximum intensity 15.5 mm/h.

As the capital of the UK, London has unique history, culture and venues that make it one of the most popular destinations for tourists and events. This brings the London transport services a big pressure. Though there are thousands of events happening every year, not all of them have a large impact on the Victoria line. In this work, we have tried to identify the events that may affect the Victoria line, which is basically judged by whether the venue of an event is close to the Victoria line. Figure 3-5 shows two events affecting the passenger flow, a concert and a football match. It can be observed that the passenger load shows a significant increase before and after the events.

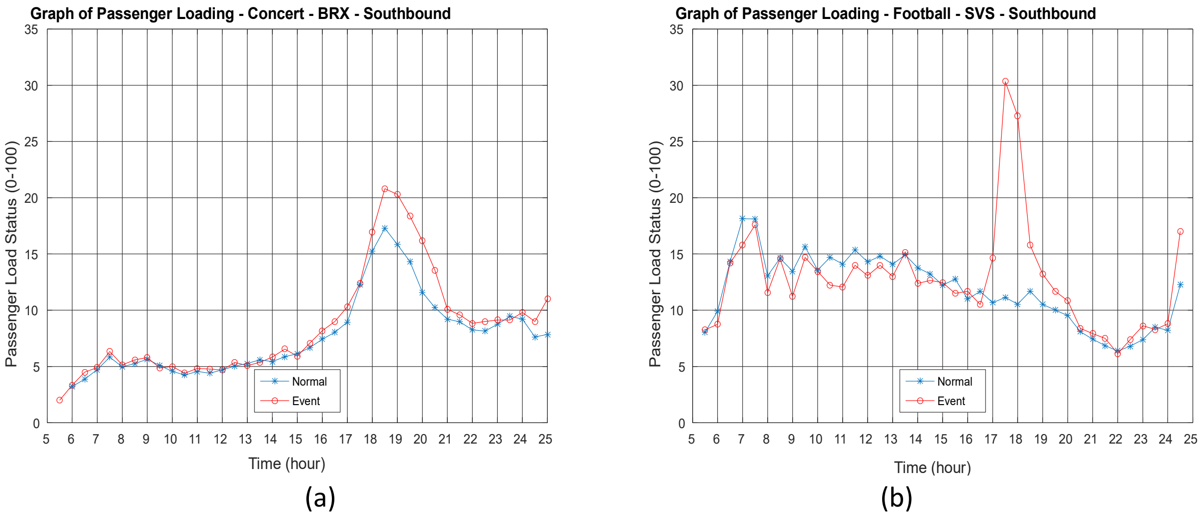


Figure 3-5: Changes on passenger flow for two events: (a) a concert close to BRX: 02 Dec 2014, event duration 19:00-23:00, increase in arriving passenger number 4859; and (b) a football match close to SVS: 11 Apr 2015, event duration 15:00-16:45, increase in leaving passenger number 9049.

There are some other factors that would also affect the number of passengers, such as accessibility of stations (Li et al. 2017b), availability of alternative means of transportation (private cars, taxis, buses, shared bicycles, etc.), and their price, quality of service and convenience (Zhang et al. 2018b, Li et al. 2019b, Zhang et al. 2019b). For a given station, their effect on the passenger number normally keeps stable, if there is no large change in transport services and environment around the station. Therefore, for prediction of the short-term passenger flow, one doesn't need to consider every single factor, but regard all of their effect has been embedded into positional and temporal factors.

4. Methodology of Data-driven Modelling

In this research, we propose a two-layer structure to achieve an accurate prediction for the passenger flow on trains, which is shown in Figure 4-1. This framework includes a global model and a number of local models. The global model can provide a general prediction for the passenger number based on temporal and positional information. The local models are able to predict the change in the passenger number due to some momentary factors, such as events and weather conditions. The aggregation of the global model output and the local models outputs provides an accurate and practical prediction for the real-time passenger number.

The two-layer modelling structure is preferable to a single model that uses all the attributes as inputs, due to two reasons: better accuracy and better maintainability. In the historical data set, the number of data samples that involve special events and special weather conditions is relatively small, which cause an imbalance between the data from normal days and those from special days. If a 'single model' is trained using all available data, it will perform well in the areas with high-density training data (normal days) but not that well in the areas with sparse training data (days with events or adverse weather). The two-layer structure can solve the problem by modelling the effect of special attributes separately, which can balance the training effort among various conditions and help improve the prediction accuracy for the special days. On the other hand, in the case that new event-related or weather-related data are collected, the 'single model' needs to be fully redeveloped and retrained, which is time-consuming. However, the two-layer structure allows only updating the relevant local model or building a new local model and integrating with other modular components.

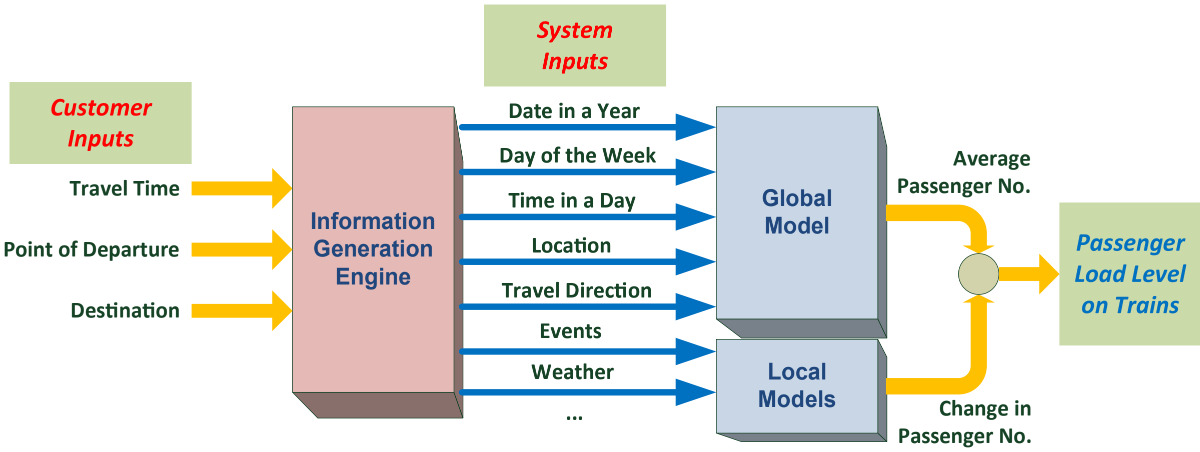


Figure 4-1: The proposed modelling framework including global model and local models.

For the generation of global and local models from data, we propose a modelling approach implementing FRBSs (Zadeh, 1973). In this approach, a heuristic algorithm reduced space searching algorithm (RSSA) (Zhang and Mahfouf, 2010) is employed to provide learning and optimisation abilities. A clustering-based data sifting strategy is proposed to select the most informative and representative data for model training, which greatly helps to improve the modelling efficiency in a case of possessing a large amount of data. This strategy is particularly important for the case of modelling the number of passengers on trains, as it considers the data from different stations and different travel directions, which are much more than the data used in modelling the passenger volume at a single station.

4.1 Introduction to fuzzy rule-based systems

FRBSs can process complex, vague and uncertain information by using interpretable linguistic rules. They were developed based on fuzzy set theories (Zadeh, 1965). A fuzzy set can be denoted as a pair (A, μ_A), where A is the set name and μ_A is the membership function. For a given value $x \in A$, $\mu_A(x)$ is its membership value with respect to A. Gaussian functions, triangular-shape functions, trapezoidal-shape functions are normally used as membership functions.

FRBSs combine fuzzy set theories with linguistic "If-Then" rules to represent systems, which makes the system models easier to be interpreted and explained. Figure 4-2 shows an example of a general FRBS, which includes four main parts: a fuzzy rule-base, fuzzifiers, a fuzzy inference engine and a defuzzifier.

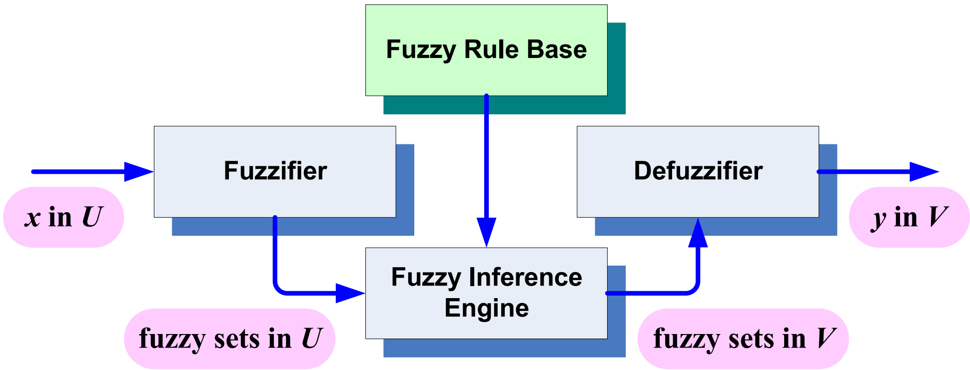


Figure 4-2: Basic configuration of fuzzy systems.

The fuzzy rule-base, also known as a knowledge-base, is the core of a fuzzy system. It consists of the following If-Then linguistic statements:

Rule$_j$: **IF** x_1 is A_1^j and x_2 is A_2^j ... and x_n is A_n^j, **THEN** y is B^j,

where $j=1,2,\ldots,m$ and m is the total number of fuzzy rules; $X=[x_1,x_1,\ldots,x_n]^T$ are input fuzzy variables and A_i^j, $i=1,2,\ldots,n$, are input fuzzy sets; y is an output fuzzy variable and B^j is an output fuzzy set.

The antecedent variables and consequent variable of fuzzy rules are all fuzzy variables. This requires a real input value to be fuzzified first to allow a further use by fuzzy rules the fuzzy reference engine. The fuzzifiers convert input real values to fuzzy set values by using membership functions. The fuzzy inference engine implements fuzzy logic principles and the applicable fuzzy rules in a reasoning process, which suggests a fuzzy set output from given fuzzy set inputs. The defuzzifier maps from the output fuzzy set to a single real value. The defuzzifier aims to specify a real-value datum point to represent the output fuzzy set appropriately. The commonly used defuzzification methods include centre average, middle of maximum, centre of gravity, etc.

There are two ways to construct FRBS models: to generate models from expert knowledge and to generate models from data. The data-driven modelling approach enables an automatic identification of the structure and parameters of FRBSs. As a primary requirement, a data-driven modelling method needs learning and optimisation abilities to initialise and refine fuzzy sets and fuzzy rules of the system. Heuristic algorithms are commonly used learning and optimisation techniques.

4.2 A fuzzy modelling approach with data sifting

In this research, we developed a data-driven modelling approach that constructs linguistic FRBSs with good accuracy and interpretability, which is named a fuzzy inference system with data sifting and structure improvement (FIS-DSSI). The FRBSs employ Gaussian functions as the membership functions of fuzzy sets, the product inference engine as the fuzzy inference engine and the centre average method as the defuzzification method (Wang 1997). In the proposed approach, a clustering technique is utilised to elicit the initial fuzzy model. A data sifting algorithm is designed to select the most compact and informative training data set. A heuristic searching and optimisation algorithm RSSA is then implemented to improve the structure and parameters of the initial model. Figure 4-3 shows the proposed modelling approach, which consists of several execution steps. It is also worth noting that the modelling approach is designed based on continuous variable domains, but it works well with discrete input variables, since one can regard a discrete input as a special case of a continuous input. The proposed fuzzy modelling method will be implemented to construct both the global and local models for passenger flow. The difference between the global model and the local model lies in different choices of input and output variables and different training data.

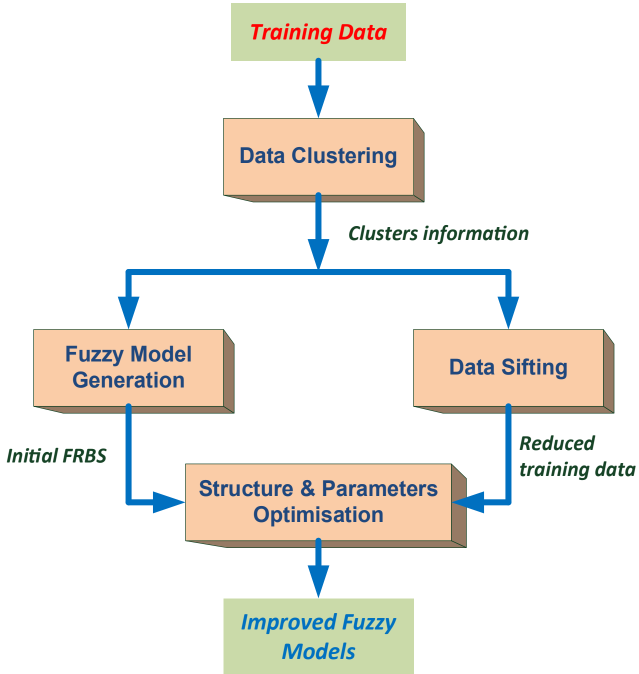


Figure 4-3: The proposed fuzzy modelling approach.

4.2.1 Clustering

Clustering is an unsupervised classification technique that groups unlabelled data into meaningful clusters. The data within one cluster share similarities and the data in different clusters are rather dissimilar. In this work, an improved clustering algorithm, the fast agglomerative complete-link clustering (Zhang and Mahfouf, 2008) was used, which was designed to greatly reduce the computational complexity of conventional hierarchical clustering algorithms. The Euclidean distance is used as the similarity measure in our work.

4.2.2 Generation of an initial fuzzy model

The clusters information is then used to build an initial FRBS. In the model, each fuzzy rule reflects a single cluster. The fuzzy sets are defined based on the position and shape of clusters. The centre of a fuzzy set's membership function is set to be the centre of the corresponding cluster. The membership function should cover the scope of the corresponding cluster.

In this work, we employ the Gaussian membership functions as follows:

$$\mu_A(x) = e^{\frac{-(x-c)^2}{\sigma^2}}, (2)$$

where $\mu_A(x)$ represents the membership degree of x belonging to A; $\sigma>0$ and it relates to the scope of the membership function; c is the membership function's centre.

As an example, we consider a general modelling problem that possesses a collection of $(n+1)$-dimensional input-output (n inputs and 1 output) training data $\{p_1, p_2, \dots, p_N\}$. $p_k=\left[x_1^k, x_2^k, \dots, x_n^k, y^k\right]^T$, where $k=1,2,\dots,N$, and N is the total number of the data samples. If a clustering process is performed, one gets a specified number (m) of clusters. If C_j represents the jth cluster, $C_j=\{p_{j1}, p_{j2}, \dots, p_{j(ND\|j)}\}$, where $j=1,2,\dots,m$, and ND_j is the data volume in the jth cluster.

The jth fuzzy rule that relates to cluster C_j is as shown in Section 4.1. For a fuzzy set A_i^j, the centre parameter of the membership function is calculated using the following equation:

$$c_i^j=\sum_{l=1}^{ND_j}\frac{x_i^{jl}}{ND_j}.(3)$$

On the other hand, a membership function needs to cover all the data that compose the related cluster. This means that, for every datum in the cluster, its membership degree needs to exceed a certain level to make sure the datum is dominated by the related fuzzy rule. Such a requirement leads to the following equation:

$$\min_l\left(\mu_{A_i^j}\left(x_i^{jl}\right)\right)=\min_l\left(e^{\frac{-\left(x_i^{jl}-c_i^j\right)^2}{\left(\sigma_i^j\right)^2}}\right)=Th,(4)$$

where $l=1,2,\dots,ND_j$. Th is a threshold value and is set to 0.25 in this paper. Equation (4) can be rewritten to the following equation to determine the parameter σ_i^j:

$$\sigma_i^j=\frac{\max\left(x_i^{jl}-c_i^j\right)}{\sqrt{-\ln(Th)}}.(5)$$

4.2.3 Data sifting

In data-driven modelling, a large number of training data may not lead to good performance and may reduce the modelling efficiency. Some data may represent the same or very similar information. Using all of such data will not be efficient. On the other hand, the available data may show different densities in different areas due to various data collection strategies. If all the data are used in modelling, the developed model will only be finely trained in certain areas and will perform badly in other areas. To avoid these situations, we designed a data sifting technique to select the most representative data with a balanced data density for model training purpose.

In this research, we utilise clusters information for training data selection. During clustering, the data samples are grouped into different clusters containing some distinctive information. This indicates that the representative training data need to be obtained from every single cluster. To balance the data distribution, one should select the same number of data from each cluster, or make the resulted data density approximately the same across different areas.

The proposed data sifting method works as the following steps:

1. If the number of data in one cluster C_j is odd, add an infinity datum to the cluster to form a new cluster $C_j=\{p_{j1}, p_{j2}, \dots, p_{j(ND\|j)}\}$, where $j=1,2,\dots,m$ and ND_j is an even number.

2. Within each cluster, randomly pair the data samples. $C_j = \{(p_{jr(1)}, p_{jr(2)}), (p_{jr(3)}, p_{jr(4)}) \dots, (p_{jr(ND_j-1)}, p_{jr(ND_j)})\}$, where $r = randperm(ND_j)$ and $randperm()$ return a row vector including a random permutation of the integers from 1 to ND_j inclusive.
3. For each cluster, do a second round paring to form $C_j^{\square} = \{(p_{jr^{\square}(1)}, p_{jr^{\square}(2)}), (p_{jr^{\square}(3)}, p_{jr^{\square}(4)}) \dots, (p_{jr^{\square}(ND_j-1)}, p_{jr^{\square}(ND_j)})\}$, where $r^{\square} = randperm(ND_j)$.
4. Combine two sets of data pairs of all clusters $\{C_1, C_2 \dots, C_m, C_1^{\square}, C_2^{\square}, \dots, C_m^{\square}\}$ and get $\{(p'_1, p'_2), (p'_3, p'_4) \dots, (p'_{(\sum_j ND_j - 1)}, p'_{(\sum_j ND_j)})\}$.
5. For each pare, calculate its distance measure using $d_k = N(p'_k, p'_{(k+1)}) = \sqrt{\sum_{i=1}^{n} (w_{Ii}^2 \times (x_i^{k+1} - x_i^k)^2 + w_O^2 \times (y^{k+1} - y^k)^2)}$, where $k = 1, 3, \dots, (\sum_j ND_j - 1)$ and $w_I = [w_{I1}, w_{I2}, \dots, w_{\square}]$ and w_o are the distance weights for input and output variables, respectively.
6. Find the pair with the minimum distance value $d_{id_{min}} = min(d_k)$, where id_{min} is the index of the pair $(p'_{id_{min}}, p'_{(id_{min}+1)})$.
7. Except for the found pair, find the datum $p'_{id_{min}}$ in another pair $(p'_{id1}, p'_{(id1+1)})$, where $p'_{id1} = p'_{id_{min}}$ or $p'_{(id1+1)} = p'_{id_{min}}$. Also, find the datum $p'_{(id_{min}+1)}$ in another pair $(p'_{id2}, p'_{(id2+1)})$, where $p'_{id2} = p'_{(id_{min}+1)}$ or $p'_{(id2+1)} = p'_{(id_{min}+1)}$.
8. If $N(p'_{id1}, p'_{(id1+1)}) < N(p'_{id2}, p'_{(id2+1)})$, remove $p'_{id_{min}}$ from clusters C_j; otherwise, remove $p'_{(id_{min}+1)}$ from clusters C_j.
9. Set $d_{id_{min}} = \infty$.
10. Repeat Steps 6 – 9 for $ceil(ND_j/10)$ times, where $ceil(x)$ rounds up x to the nearest integer. After the iterations, about 10% of the data in the current clusters are removed.
11. If the total number of data in all clusters is smaller than a predefined value NR, terminate the process; otherwise, return to Step 1.

The distance weights w_I and w_o will affect the compactness of clusters and will thus affect the elicited FRBS's sensitivity to a certain input variable. Generally, a large value of a weight suggests that the variable of the corresponding dimension is more important than another variable with a small weight value. For the modelling purpose, it is appropriate to use a large value for the output weight.

4.2.4 Structure improvement

The structure and parameters optimisation process starts after the completion of initial model generation and data sifting. The optimisation process repeats whenever new parameters are generated and stops until a termination criterion is met. The optimisation process is carried out based on the reduced training data, which can help reduce the computational burden and increase the training speed.

In this work, the structure and interpretability of models are improved using the following operations: combination of similar fuzzy sets, elimination of redundant fuzzy sets, combination of similar fuzzy rules and elimination of redundant fuzzy rules. Through the reduction of fuzzy sets and fuzzy rules, the total number of parameters is reduced and the complexity of the model is reduced.

The similarity measure between fuzzy sets is defined as follows:

$$S\left(A_i^k, A_i^l\right)=\frac{1}{1+\sqrt{\left(c_i^k-c_i^l\right)^2+\left(\sigma_i^k-\sigma_i^l\right)^2}} \quad (6)$$

A threshold Th_1 is used to control the combination of similar fuzzy sets. If $S\left(A_i^k, A_i^l\right)>Th_1$, which means the fuzzy sets A_i^k and A_i^l are highly overlapped. In this case, they will be combined to a single set A_i^{new}. A fast way to calculate its parameters is that $c_i^{new}=\left(c_i^k+c_i^l\right)/2$ and $\sigma_i^{new}=\left(\sigma_i^k+\sigma_i^l\right)/2$. One can also apply Equations (3) and (5) to calculate c_i^{new} and σ_i^{new} by merging samples in relevant clusters.

Redundancy of a fuzzy set A_i^k is identified by calculating the similarity $S\left(A_i^k, U\right)$, where U is the universal set and $\mu_U(x)=1, \forall\ x$. If a fuzzy set is very similar to the universal set, the membership degree of any given value according to the set is always close to 1. Such a set does not provide any help in differentiating different situations and is considered to be redundant. A threshold Th_2 is used to control the elimination of redundant fuzzy sets. If $S\left(A_i^k, U\right)>Th_4$, A_i^k is deemed to be redundant and will be removed. In the case that a redundant set is removed, the antecedent conditions of fuzzy rules that include the set should also be removed. This will result in some 'shorter' fuzzy rules that do not contain all the input variables.

The similarity between two fuzzy rules is determined by the similarity of the antecedent terms of these two rules. The similarity measure is designed as follows:

$$S_R\left(R_k, R_l\right)=\prod_{i=1}^{n} S\left(A_i^k, A_i^l\right) \quad (7)$$

where $A_1^k, A_2^k, \ldots, A_n^k$ are the fuzzy sets in the preconditions of the kth fuzzy rule R_k; $A_1^l, A_2^l, \ldots, A_n^l$ are the fuzzy sets in the preconditions of the lth fuzzy rule R_l. Once S_R exceeds a threshold value Th_3, R_k and R_l will be combined to form a new fuzzy rule R_{new}. The pairs of the fuzzy sets involved in both the antecedent and consequent terms of R_k and R_l will be merged, following the method introduced before.

To evaluate the redundancy of a linguistic rule, we employed an evaluation measure *confidence* (Ishibuchi *et al*. 2001). If C is a set with N training data $p_k=\left[x^k, y^k\right]^T=\left[x_1^k, x_2^k, \ldots, x_n^k, y^k\right]^T$, where $k=1,2,\ldots,N$. The *confidence* of a fuzzy rule was designed as the following equation:

$$conf(A\rightarrow B)=\frac{|C(A)\cap C(B)|}{|C(A)|}=\frac{\sum_{k=1}^{N}\left(\mu_A\left(x^k\right)\times\mu_B\left(y^k\right)\right)}{\sum_{k=1}^{N}\mu_A\left(x^k\right)} \quad (8)$$

where $|C(A)|$ represents how many data samples that are compatible with antecedent conditions; $|C(A)\cap C(B)|$ represents the approximate number of data samples, which are consistent with both the antecedent and consequent conditions; $\mu_A\left(x^k\right)$ is the compatibility level for the input vector x^k with respect to terms $A=[A_1, A_2, \dots, A_n]$; and $\mu_B(y^k)$ is the compatibility level for the output y^k to a consequent term B. $\mu_A(x^k)$ was usually defined by a product operator as follows:

$$\mu_A(x^k)=\mu_{A_1}(x_1^k)\times\mu_{A_2}(x_2^k)\times\dots\times\mu_{A_n}(x_n^k)\,(9).$$

Confidence measures the proportion of data that are compatible with a certain fuzzy rule. If the value is low, it means the fuzzy rule is poorly supported by data and a large number of observations may be against it. Normally, the fuzzy rules generated using a clustering-based method have good *confidence* values, while a grid-partitioning-based method may generate some redundant rules with low *confidence* values. In this work, if the index value of a certain rule is less than a threshold Th_4, the rule is considered redundant and will be removed.

4.2.5 Structure and parameters optimisation

The structure improvement mechanism introduced in Section 4.2.4 can only improve the simplicity (or interpretability) of a FRBS and may cause a decrease in modelling accuracy. Therefore, it must cooperate with a parameter optimisation mechanism to achieve a balanced improvement between accuracy and interpretability. We consider two aspects in the improvement of an initial FRBS: accuracy and interpretability. The Root Mean Square Error ($RMSE$) is used to evaluate the accuracy performance and is defined as follows:

$$RMSE=\sqrt{\frac{\sum_{k=1}^{N}\left(y_m^k-y_p^k\right)^2}{N}}\,(10)$$

where y_m^k represents a measured output, y_p^k represents a predicted output, $k=1,2,\dots,N$ and N is the total data number. The interpretability relates to the average length of fuzzy rules ($Lrule$), the quantity of fuzzy rules ($Nrule$) and the quantity of fuzzy sets ($Nset$).

To normalise the measures used to assess accuracy and interpretability and make them comparable in scale, we design an objective function for optimisation as follows:

$$w_1\frac{RMSE}{RMSE_I}+w_2\frac{Lrule}{Lrule_I}+w_3\frac{Nrule}{Nrule_I}+w_4\frac{Nset}{Nset_I}\,(11)$$

where $RMSE_I$ is the $RMSE$ of the initial FRBS; $Lrule_I$, $Nrule_I$ and $Nset_I$ are the average rule length, fuzzy rules number and fuzzy sets number of the initial FRBS, respectively.

A heuristic optimisation algorithm is utilised to tackle the objective function to improve both the structure and parameters of fuzzy systems. The decision variables include the parameters of all fuzzy sets and four thresholds Th_1 - Th_4, which control the degree of interpretability improvement as introduced in Section 4.2.4. Figure 4-5 shows a flow diagram of the designed mechanism for structure and parameters optimisation.

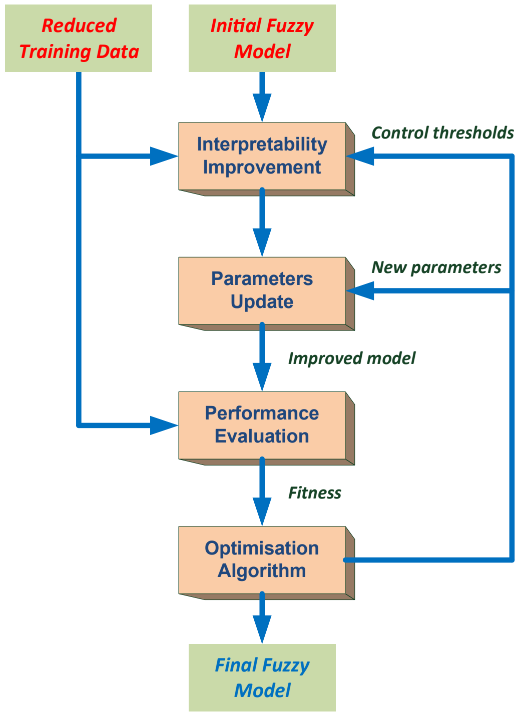


Figure 4-5: The proposed mechanism for structure and parameters optimisation.

In this work, we employed a fast search and optimisation method, reduced space searching algorithm (RSSA) (Zhang and Mahfouf 2007, 2010). It was inspired by a human behaviour when searching for a solution in daily life. Different with other heuristic algorithms, RSSA tries to transform and shift the search space to locate the optimal sub-space, while most other optimisation techniques focus on producing new candidate solutions with various mechanisms. RSSA was shown to be a successful algorithm that can find the global optimum rapidly.

5. Predictive Models of Passenger Flow on Trains

The designed fuzzy modelling method FIS-DSSI has been implemented to build both the global and local models for passenger flow of London Underground trains. The difference between the global model and the local model lies in different choices of input and output variables and different training data. In the modelling experiments, the initial FRBS model was constructed with 50 fuzzy rules and 50 fuzzy sets for each dimension, which were later reduced through the optimisation process. The configurations of the RSSA algorithm were set the same as the recommendations in (Zhang & Mahfouf, 2010). The maximum function evaluation number is 10,000. Each experiment was run multiple times and the results showed very good repeatability and consistency.

5.1 Global model

To predict the passenger flow in an ordinary condition, four attributes were used as the model inputs, which are time in a day, day of the week, station ID and travel direction. For day of the week, discrete values 1 to 7 were assigned to represent Monday to Sunday, respectively. For station ID, discrete values 0 to 15 were assigned to represent the stations from the south terminus to the north terminus, respectively. It means the station ID for Brixton is 0 and the station ID for Walthamstow Central is 15. For travel direction, 0 represents northbound trains and 1 represents southbound trains. The model output is the average passenger load of all the trains approaching a particular station within a 30-min time interval. It is worth noting that the passenger load approaching one station is the same with the passenger load departing from the previous station. Since the model is developed to predict the passenger load arriving at every station, the information about departing passenger load can also be obtained from the arriving prediction of the next station. In the experiments, 10685 data were used in training and other 7459 data were used in testing. After the operation of the proposed data sifting mechanism (see Section 4.2.3), 3326 data were obtained as the representatives for the whole training data set and they have shown to work very well in the generation of the global model. Figure 5-1 shows the performance of the generated model in the prediction of both the training and testing data sets. The RMSE of this model is 1.5457 for training and 1.8358 for testing.

To validate the model in prediction, it was further employed to predict the passenger flow of certain trains. Figure 5-2 shows the forecasting examples for both northbound and southbound trains, at the stations Oxford Circus and Stockwell, on Wednesday and Saturday. From the figure, it is observed that the developed model can forecast the passenger load dynamically and precisely. The model is reckoned to be robust, as it tends to produce moderate and smooth predictions and ignore the noises involved.

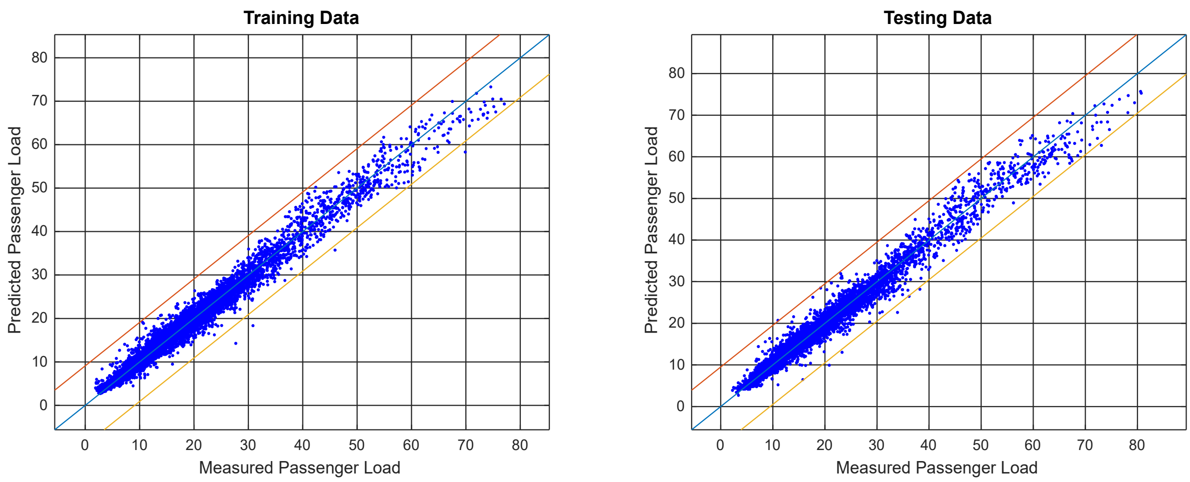


Figure 5-1: Measured passenger load versus predicted passenger load using the developed global model (with ±10% error bands).

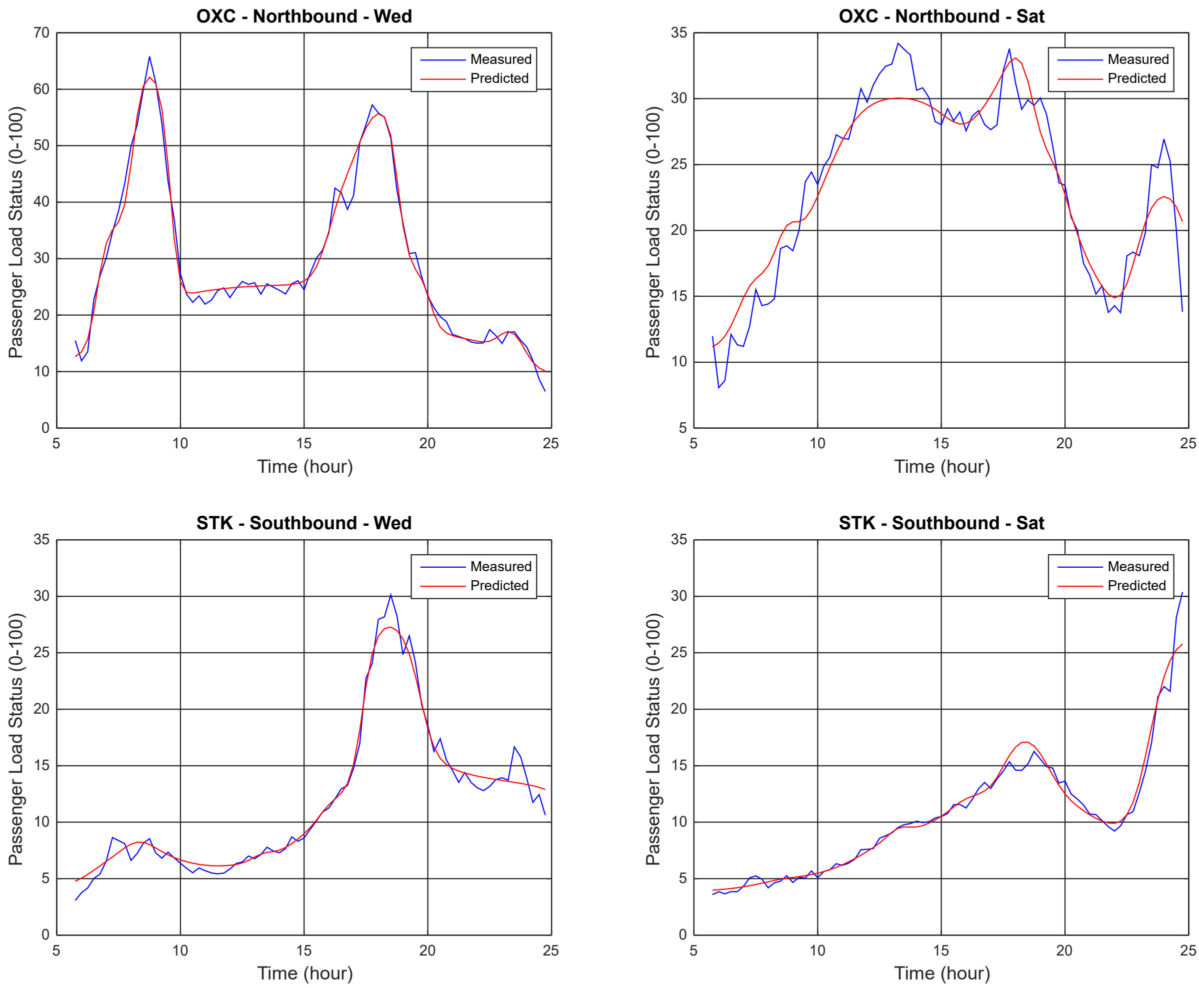
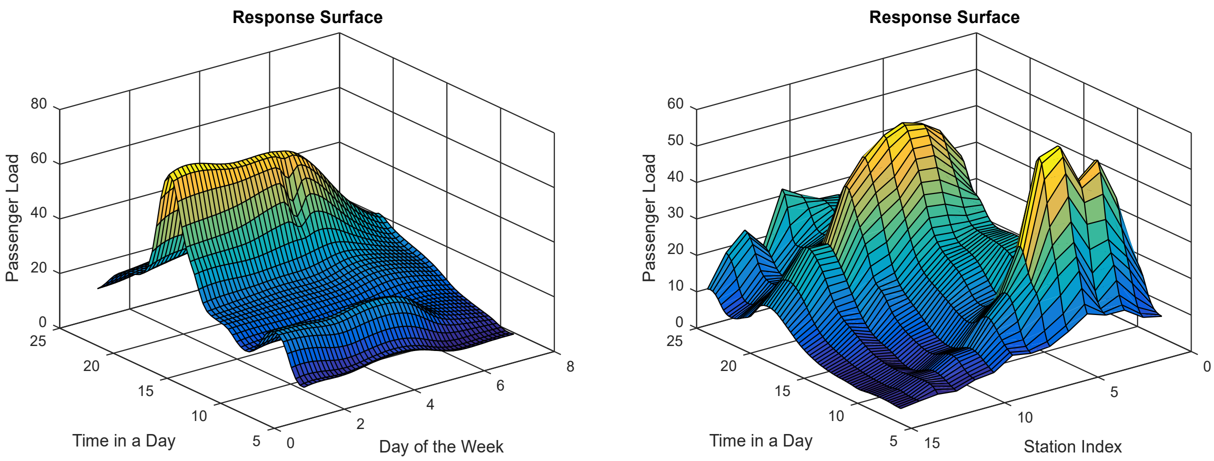


Figure 5-2: Prediction of the passenger load on northbound and southbound trains when approaching Oxford Circus and Stockwell.

Figure 5-3 includes the three-dimensional response surfaces of the generated model. From Figure 5-3(a), it is observed that there is a distinct afternoon peak for taking the northbound trains at Euston; and the maximum passenger load for this afternoon peak is slowly going down from Monday to Friday, and then largely decreasing on Saturday and Sunday. Figure 5-3(b) demonstrates that the model can predict different passenger load patterns for different stations, for example, the south stations have a high morning peak while some middle stations have a high afternoon peak. This is consistent to the data observation done in Section 2 and it reflects the fact that some commuters take the northbound trains to work in the morning and some other commuters travel from the city centre to north to go back home in the afternoon. The smooth response surfaces also demonstrate the good nonlinear mapping and generalisation abilities of the proposed modelling approach.



(a) (b)

Figure 5-3: 3-D response surfaces of the global model: (a) for northbound trains at Euston and (b) for northbound trains on Friday.

Figure 5-4 illustrates 5 fuzzy rules of the developed model. Such fuzzy rules may also be described as approximate linguistic rules, which employ some linguistic hedges (Zadeh, 1972), such as:

R_1: IF Station Index is *small* AND Travel Direction is '1' AND Day of the Week is *small* AND Time in a Day is *more or less medium*, THEN Passenger Load is *small*.

R_3: IF Station Index is *small medium* AND Travel Direction is '1' AND Day of the Week is *more or less small medium* AND Time in a Day is *medium large*, THEN Passenger Load is *large*.

Through investigating the linguistic rules, one may obtain more knowledge about the passenger flow on trains and the potential relationships between inputs and output.

Table 5-1 compares the performance of the proposed FIS-DSSI method with some parametric and non-parametric models. ANN-BP represents a feedforward neural network with a training method of Bayesian regularization backpropagation (Dan Foresee and Hagan 1997). It has a single hidden layer with 50 sigmoid neurons. The adaptive neuro-fuzzy inference system (ANFIS) model is a Sugeno-type fuzzy inference system with a combination of the least-squares and back-propagation gradient descent methods in training (Jang 1993). The ANFIS model generated in experiments has 53 fuzzy rules. The ANN-BP and ANFIS models used the same model inputs and training and testing data as the FIS-DSSI model. It can be observed that the newly developed model outperforms these widely used neural network and fuzzy models. A single AR or ARIMA model is only developed for a single station for either northbound or southbound trains, where Table 5-1 shows the example of such models for southbound trains at station OXC. In the experiments, the order 2 AR model and the ARIMA (2, 0, 2) (Ding et al. 2018) model were employed. The results in Table 5-1 show the accuracy of prediction for future one step (30 minutes). The time-series models are suitable to predict for the near future based on the current and near past observations, while the non-parametric models do not have such a limitation.

R_1 R_2 R_3 R_4 R_5

IF Station Index is

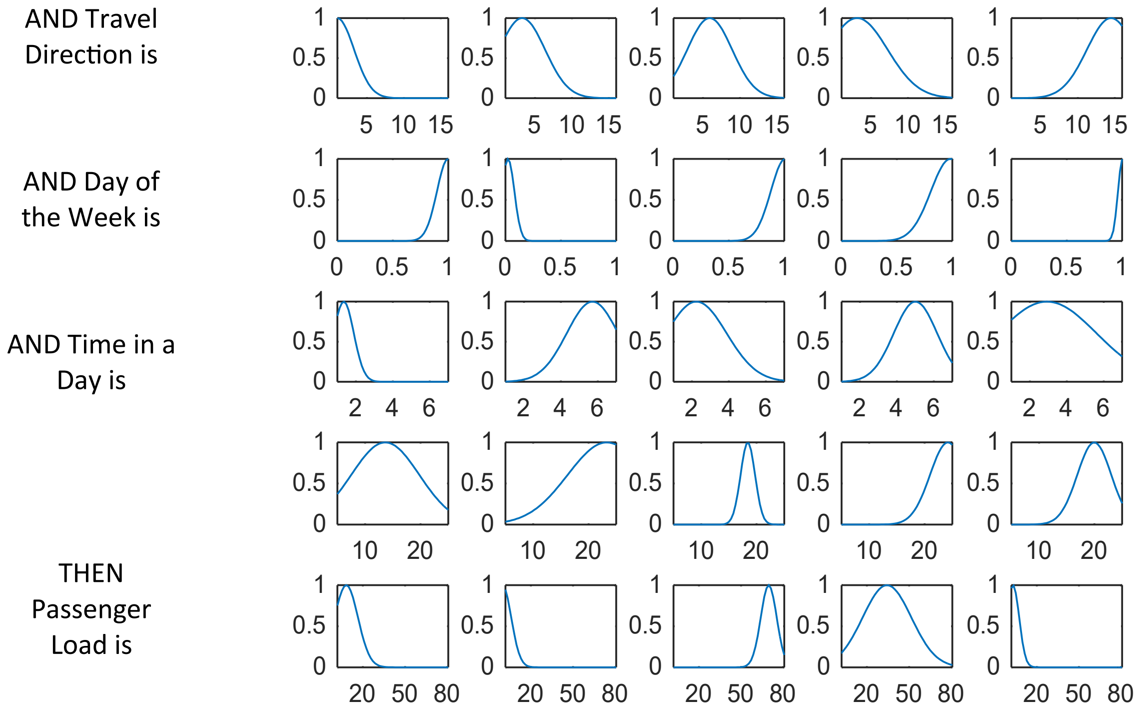


Figure 5-4. Examples of fuzzy rules of the global passenger flow model.

Table 5-1: Comparison of different methods in modelling the passenger flow.

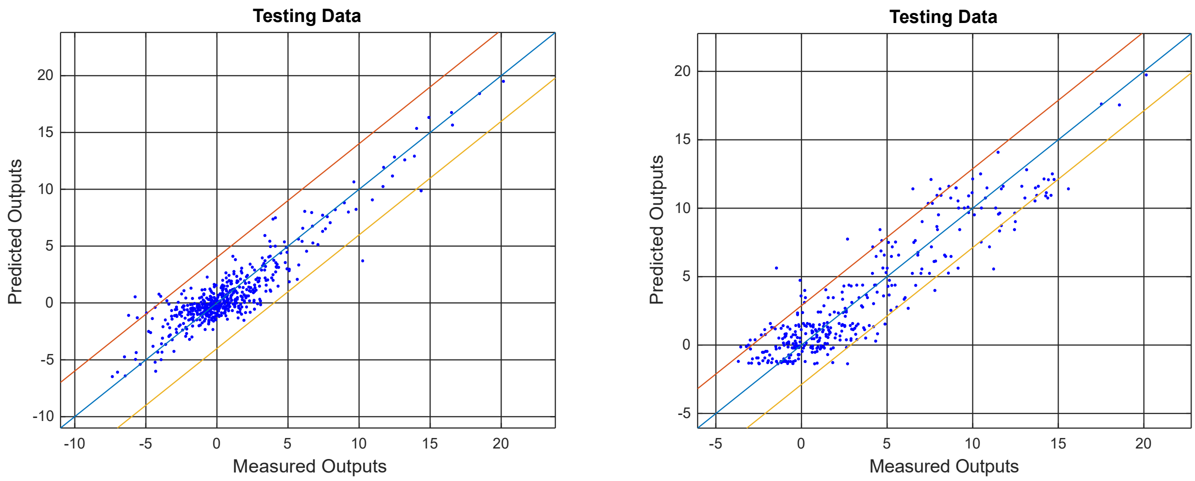
Model	Training		Testing	
	RMSE	*r*	*RMSE*	*r*
AR*	3.5022±0	0.9609±0	3.8235±0	0.9591±0
ARIMA*	3.3425±0	0.9630±0	3.6175±0	0.9619±0
ANN-BP	2.4928±0.0971	0.9790±0.0018	2.6187±0.0951	0.9767±0.0015
ANFIS	3.1483±0	0.9658±0	3.2561±0	0.9646±0
Proposed FIS-DSSI	1.5457±0.0311	0.9880±0.0032	1.8358±0.1491	0.9883±0.0008

*Models for passenger flow of southbound trains at OXC.

5.2 Local models

Two local models were developed, which can forecast the change in passenger load due to some temporary factors, i.e. weather conditions and events. It should be noted that the service disruption and its spatiotemporal effects (Li et al. 2020) would also affect the passenger flow. If the relevant data are available, one can construct another local model to predict the change of passenger flow caused by service disruption and its spatiotemporal effect. In the weather local model of the current case study, the only weather condition considered is heavy rains. Some other adverse weather conditions, such as snows and heavy winds, were not included into the current model due to the lack of the relevant data.

For the weather local model, the input variables include time in a day, day of the week, station ID, travel direction, rainfall intensity and estimated water volume of a rainy day. The assignment of the discrete variables is the same with that of the global model. The output variable is the passenger load change on the trains of the Victoria line. In the model construction, 1032 data samples were used in training and 518 data samples were used in testing. The RMSEs of this model for training and testing are 1.4295 and 1.6734, respectively. Figure 5-5(a) gives the predictive performance of the generated local model working on testing data.



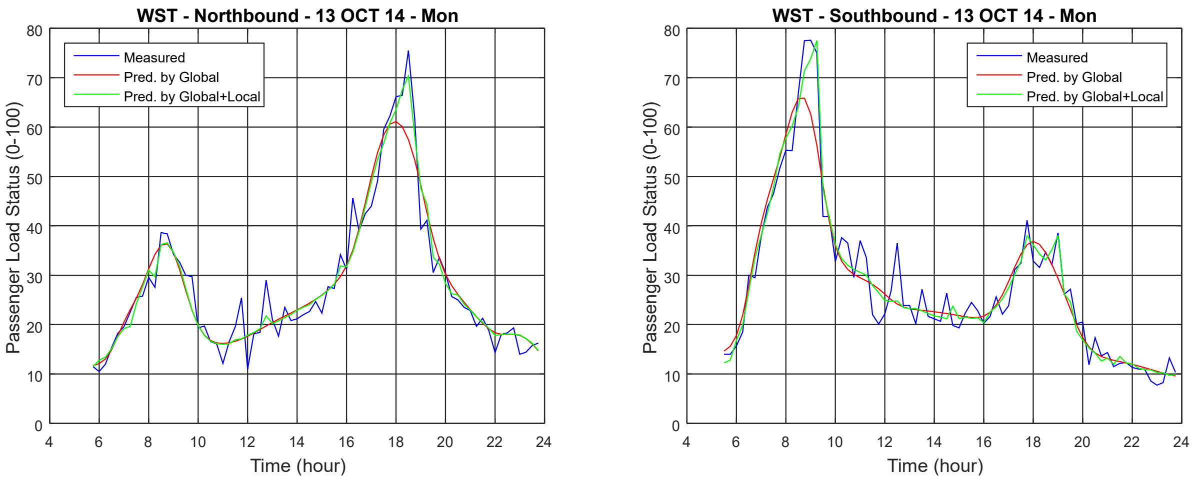
(a) (b)

Figure 5-5: Measured versus predicted outputs using developed local models (with ±10% error bands): (a) the weather local model and (b) the event local model.

To validate the model in prediction, the weather local model was employed to predict the variation of the passenger load in a rainy day, 13th October 2014, Monday. As shown in Figure 5-6, the local model helps adjust the prediction provided by the global model. The aggregation of the global model output and the local model output provides an accurate prediction for the true passenger load level.

In this study, only the events that regularly happen and have a large impact on the passenger number of the Victoria Line were included in the event local model. We can identify two types of events with different patterns in affecting passenger number. One type of events are ‘time-concentrated’ events, such as football matches and concerts, where people are expected to attend the whole period of an event and thus they normally take trains before and after the event. The other type of events are ‘time- flexible’ events, such as festive fairs, where people can join at any time during an event which results in an increase in the train passenger number during the whole period of the event.

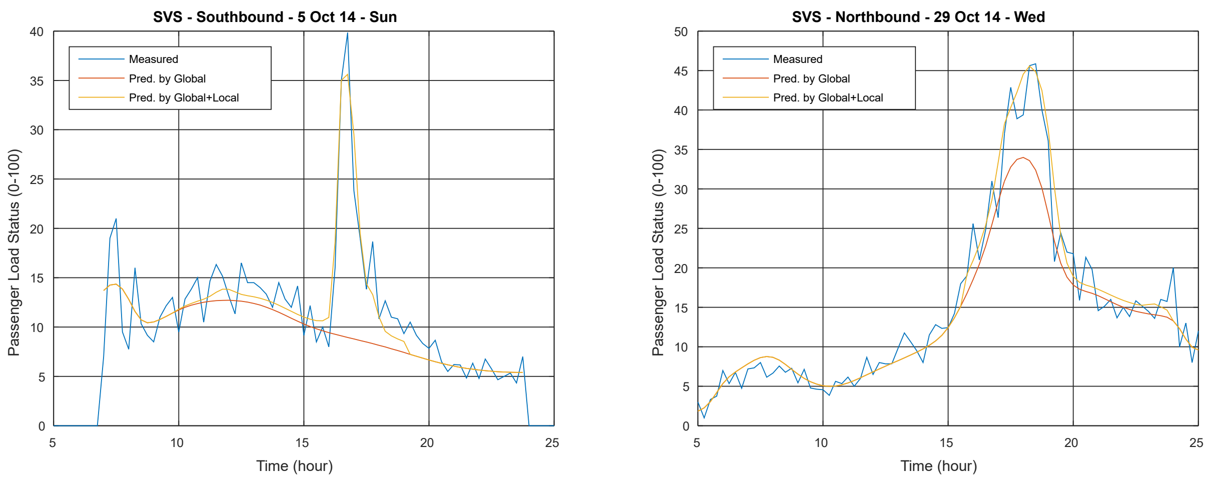
For the event local model, the input variables include time in a day, day of the week, station ID, travel direction, event ID, start time of event, end time of event and estimated attendance. The assignment of the discrete variables is the same with that of the global model. For event ID, the value assignment can be found in Table A-2 in Appendix. The output variable is the passenger load change on the trains of the Victoria line. In the model construction, 714 data samples were utilised in training and 357 data samples were utilised in testing. The RMSEs of this model for training and testing are 1.8774 and 2.0786, respectively. Figure 5-5(b) gives the predictive performance of the generated local model working on testing data.



(a) (b)

Figure 5-6: Prediction of passenger load on trains in a rainy day, 13th October 2014, Monday: (a) northbound trains approaching Warren Street and (b) southbound trains leaving Warren Street.

Figure 5-7 shows the examples of using the global model plus the local event model to predict the passenger load for two football-match days. The first day is 5th October 2014, Sunday, and an England Premier League football match (Tottenham Hotspur vs. Southampton) was held in White Hart Lane at 14:00 – 15:50. The second day is 29th October 2014, Wednesday, and an England League Cup match (Tottenham Hotspur vs. Brighton & Hove Albion) was held in White Hart Lane at 19:45 – 21:30. From the figure, we can observe a good consistency between the predicted values and the measured real data. This means that the local models help adjust the prediction by the global model to achieve more accurate prediction for the passenger flow.



(a) (b)

Figure 5-7: Prediction of passenger load on trains on two football-match days: (a) southbound trains leaving Seven Sisters on 5th October 2014, Sunday and (b) northbound trains approaching Seven Sisters on 29th October 2014, Wednesday.

6. Conclusion

In this paper, we have proposed a hierarchical modelling framework for traffic/passenger flow prediction. It includes two layers of models, where a global model used to predict the passenger flow under ordinary circumstances and a number of local models used to predict the changes in passenger flow due to specific factors, such as events, weather and service disruption. A new data sifting method has been proposed to select the most informative and representative data for model training, which greatly reduces the computational complexity and improves the modelling efficiency. The proposed method has been successfully applied to the problem of forecasting passenger flow of London Underground trains. It was shown that the aggregation of the global model output and the local models outputs provides an accurate and practical prediction for the real-time passenger flow.

The case study shows that the passenger flow greatly varies in different locations at different times. On weekdays, the passenger flow in the morning peak and the afternoon peak could be 2-3 times of that in off-peak hours. The passenger flow at the stations in the city centre could be 2-3 times of that at the stations close to the termini. It is also noticed that there is an increase in passenger flow during the late nights on Friday and Saturday. The rail operators and station administrators should allocate appropriate train capacity, personnel, and service and infrastructural facilities to reflect the spatiotemporal distribution of passenger flow. The passenger flow analysis and prediction also helps formulate maintenance schedule by finding less busy time, helps prepare and evaluate emergency response plans by providing possible passenger number, and helps prioritise safety and security inspection by providing spatiotemporal distribution of passenger flow. Considering the high passenger demand in the weekday peak hours, more services may be introduced to the timetable to avoid overcrowded scenarios. For the stations with low passenger flow, their attraction and the radiation accessibility (Li et al. 2017b.) should be evaluated. If needed, the accessibility can be improved by increasing the connection buses, improving the nearby walking environment and providing enough shared bicycles.

It shows that the impact of weather on the subway (Metro) passenger flow is not very large. Under some adverse weather conditions, such as heavy rains, some people may cancel or delay their inessential trips; on the other hand, some people tend to shift travel modes from walking, cycling and bus to subway. Considering a large passenger number under bad operating conditions, such as slippery surfaces in stations and trains, some extra risk is introduced. Therefore the risk mitigation strategy should be carefully designed and followed. Large events greatly affect the passenger flow in a short time window. If the surge of passenger number happens in peak hours, it can easily lead to overcrowded and unsafe scenarios. The transport regulators and policymakers should take this into account when scheduling and approving big sport and culture events; and the transport operators may introduce extra subway and/or other transport services, provide more personnel and resources to ease the short-term pressure relating to events.

There are still some limitations in this work. For instance, the current case study has analysed the passenger flow data from one train line but has not addressed the interrelationship between different lines; the current case study has not investigated how service disruption and its spatiotemporal effects would affect the passenger flow. Some future work can be carried out for further improvement and exploitation of this research. First, the current data-driven modelling work can be extended from one train line to a regional or national network, where the service disruption and its spatiotemporal effects will be studied. The interrelationship between different lines can be studied using a networked modelling paradigm. Second, the developed passenger flow models can be merged with online journey planning tools, which can provide detailed crowding prediction in advance of travel to help passengers make better decisions in their journey planning. Multi-criteria decision making (MCDM) techniques can be employed to consider the passengers' personalised preferences and suggest the best decision. Last, the developed models can be combined with multi-objective optimisation techniques to develop a decision support tool, which is used to plan the distribution of rolling stock to achieve the optimal use of train capacity.

Appendix

Table A-1: Station abbreviations.

Abbreviations	Station names	Abbreviations	Station names
WAL	Walthamstow Central	WST	Warren Street
BHR	Blackhorse Road	OXC	Oxford Circus
TTH	Tottenham Hale	GPK	Green Park
SVS	Seven Sisters	VIC	Victoria
FPK	Finsbury Park	PIM	Pimlico
HBY	Highbury & Islington	VUX	Vauxhall
KXX	King's Cross St. Pancras	STK	Stockwell
EUS	Euston	BRX	Brixton

Table A-2: Value assignment for event ID

Event	Event ID value
Football match	1
Concert and show	2
Festive fairs	3
Market	4

Acknowledgement

This work was supported by the Rail Safety and Standards Board (RSSB) and the Rail Delivery Group (RDG) [grant number COF-ICE-02]. The authors also wish to thank Transport for London (TfL) for providing the data needed for the research.

References:

Arana, P., Cabezudo, S., Peñalba, M., 2014. Influence of weather conditions on transit ridership: a statistical study using data from Smartcards. Transportation Research Part A 59, 1–12.

Bao, Y., Xiong, T., Hu, Z., 2012. Forecasting air passenger traffic by support vector machines with ensemble empirical mode decomposition and slope-based method. Discrete Dynamics in Nature and Society 2012, Article ID 431512, http://dx.doi.org/10.1155/2012/431512.

Chan, K.Y., Dillon, T.S., Singh, J., Chang, E., 2012. Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg–Marquardt algorithm. IEEE Transactions on Intelligent Transportation Systems 13(2), 644–654.

Chen, C., Wang, Y., Li, L., Hu, J., Zhang, Z., 2012. The retrieval of intra-day trend and its influence on traffic prediction. Transportation Research Part C 22, 103–118.

Chen, E., Ye, Z., Wang, C., Xu, M., 2020. Subway passenger flow prediction for special events using smart card data. IEEE Transactions on Intelligent Transportation Systems 21(3), 1109–1120.

Chen, H.K., Wu, C.J., 2012. Travel time prediction using empirical mode decomposition and gray theory. Transportation Research Record: Journal of the Transportation Research Board 2324, 11–19.

Chen, Q., Wen, D., Li, X., Chen, D., Lv, H., Zhang, J., Gao, P., 2019. Empirical mode decomposition based long short-term memory neural network forecasting model for the short-term metro passenger flow. PLoS ONE 14(9), e0222365, https://doi.org/10.1371/journal.pone.0222365.

Clark, S., 2003. Traffic prediction using multivariate nonparametric regression. Journal of Transportation Engineering 129(2), 161–168.

Dan Foresee, F., Hagan, M. T., 1997. Gauss-Newton approximation to Bayesian learning. Proceedings of International Conference on Neural Networks (ICNN'97), pp. 1930–1935.

Ding, C., Duan, J., Zhang, Y., Wu, X., Yu, G., 2018 Using an ARIMA-GARCH modeling approach to improve subway short-term ridership forecasting accounting for dynamic volatility. IEEE Transactions on Intelligent Transportation Systems 19(4), 1054–1064.

Ding, C., Wang, D., Ma, X., Li, H., 2016. Predicting short-term subway ridership and prioritizing its influential factors using gradient boosting decision trees. Sustainability 8(11), 1100, https://doi.org/10.3390/su8111100.

Dou, F., Jia, L., Wang, L., Xu, J., Huang, Y., 2014. Fuzzy temporal logic based railway passenger flow forecast model. Computational Intelligence and Neuroscience 2014, Article ID 950371, https://doi.org/10.1155/2014/950371.

Faraway, J., Chatfield, C., 1998. Time series forecasting with neural networks: a comparative study using the airline data. Applied Statistics 47(2), 231–250.

Gong, Y., Li, Z., Zhang, J., Liu, W., Zheng, Y., Kirsch, C., 2018. Network-wide crowd flow prediction of Sydney trains via customized online non-negative matrix factorization. Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp. 1243–1252.

Guo, J., Xie, Z., Qin, Y., Jia, L., Wang, Y., 2019. Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM. IEEE Access 7, 42946–42955.

Hänseler, F.S., van den Heuvel, J.P.A., Cats, O., Daamen, W., Hoogendoorn, S.P., 2020. A passenger-pedestrian model to assess platform and train usage from automated data. Transportation Research Part A 132, 948–968.

He, G., Ma, S., 2002. A study on the short-term prediction of traffic volume based on wavelet analysis. Proceedings of IEEE 5th International Conference on Intelligent Transportation Systems, pp. 731–735.

He, W., Zhong, C., Sotelo, M.A., Chu, X., Liu, X., Li, Z., 2019. Short-term vessel traffic flow forecasting by using an improved Kalman model. Cluster Computing 22, 7907–7916.

Hörcher, D., Graham, D.J., Anderson, R.J., 2017. Crowding cost estimation with large scale smart card and vehicle location data. Transportation Research Part B 95, 105–125.

Hsu, C.I., Wen, Y.H., 1998. Improved grey prediction models for the trans-pacific air passenger market. Transportation Planning and Technology 22(2), 87–107.

Ishak, S., Kotha, P., Alecsandru, C., 2003. Optimization of dynamic neural network performance for short-term traffic prediction. Transportation Research Record: Journal of the Transportation Research Board 1836, 45–56.

Ishibuchi, H., Yamamoto, T., Nakashima, T., 2001. Fuzzy data mining: effect of fuzzy discretization. Proceedings of the First IEEE International Conference on Data Mining, pp. 241–248.

Jiang, X., Adeli, H., Asce, H.M., 2005. Dynamic wavelet neural network model for traffic flow forecasting. Journal of Transportation Engineering 131, 771–779.

Jiang, X., Zhang, L., Chen, M.X., 2014. Short-term forecasting of high-speed rail demand: a hybrid approach combining ensemble empirical mode decomposition and gray support vector machine with real-world applications in China. Transportation Research Part C 44, 110–127.

Jiao, P., Li, R., Sun, T., Hou, Z., Ibrahim, A., 2016. Three revised Kalman filtering models for short-term rail transit passenger flow prediction. Mathematical Problems in Engineering 2016, Article ID 9717582, https://doi.org/10.1155/2016/9717582.

Karlaftis, M.G., Vlahogianni, E.I., 2011. Statistical methods versus neural networks in transportation research: differences, similarities and some insights. Transportation Research Part C 19, 387–399.

Kim, K.I., Lee, K.M., 2018. Deep learning-based caution area traffic prediction with automatic identification system sensor data. Sensors (Basel) 18(9), 3172, https://doi.org/10.3390/s18093172.

Kim, K.M., Hong, S.-P., Ko, S.-J., Kim, D., 2015. Does crowding affect the path choice of metro passengers? Transportation Research Part A 77, 292–304.

Kumar, S.V., Vanajakshi, L., 2015. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. European Transport Research Review 7(3), 1–9.

Lee, S., Fambro, D.B., 1999. Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting. Transportation Research Record: Journal of the Transportation Research Board 1678(1), 179–188.

Leng, B., Zeng, J., Xiong, Z., Lv, W., Wan, Y., 2013. Probability tree based passenger flow prediction and its application to the Beijing subway system. Frontiers of Computer Science 7(2), 195–203.

Jang, J.-S.R., 1993. ANFIS: adaptive-network-based fuzzy inference systems. IEEE Transactions on Systems, Man, and Cybernetics 23 (3), 665-685.

Jia, F., Li, H., Jiang, X., Xu, X., 2019. Deep learning-based hybrid model for short-term subway passenger flow prediction using automatic fare collection data. IET Intelligent Transport Systems 13(11), 1708-1716.

Jing, Z., Yin, X., 2020. Neural network-based prediction model for passenger flow in a large passenger station: an exploratory study. IEEE Access 8, 36876–36884.

Li, H., Wang, Y., Xu, X., Qin, L., Zhang, H., 2019a. Short-term passenger flow prediction under passenger flow control using a dynamic radial basis function network. Applied Soft Computing 83, Article ID 431512, https://doi.org/10.1016/j.asoc.2019.105620.

Li, L., Chen, X., Zhang, L., 2014. Multimodel ensemble for freeway traffic state estimations. IEEE Transactions on Intelligent Transportation Systems 15(3), 1323–1336.

Li, L., Ren, H., Zhao, S., Duan, Z., Zhang, Y., Zhang, A., 2017b. Two dimensional accessibility analysis of metro stations in Xi'an, China. Transportation Research Part A 106, 414–426.

Li, L., Wang, Y., Zhong, G., Zhang, J., Ran, B., 2018. Short-to-medium term passenger flow forecasting for metro stations using a hybrid model. KSCE Journal of Civil Engineering 22(5), 1937–1945.

Li, R., Jiang, C., Zhu, F., Chen, X., 2016. Traffic flow data forecasting based on interval type-2 fuzzy sets theory, IEEE/CAA Journal of Automatica Sinica 3(2), 141–148.

Li, T., Rong, L., 2020. A comprehensive method for the robustness assessment of high-speed rail network with operation data: A case in China. Transportation Research Part A 132, 666–681.

Li, T., Rong, L., Yan, K., 2019b. Vulnerability analysis and critical area identification of public transport system: A case of high-speed rail and air transport coupling system in China. Transportation Research Part A 127, 55–70.

Li, Y., Wang, X., Sun, S., Ma, X., Lu, G., 2017a. Forecasting short-term subway passenger flow under special events scenarios using multiscale radial basis function networks. Transportation Research Part C 77, 306–328.

Li, Z.-C., Sheng, S., 2016. Forecasting passenger travel demand for air and high-speed rail integration service: A case study of Beijing-Guangzhou corridor, China. Transportation Research Part A 94, 397–410.

Ling, X., Huang, Z., Wang, C., Zhang, F., Wang, P., 2018. Predicting subway passenger flows under different traffic conditions. PLoS ONE 13(8), e0202707, https://doi.org/10.1371/journal.pone.0202707.

Lippi, M., Bertini, M., Frasconi, P., 2013. Short-term traffic flow forecasting: an experimental comparison of time-series analysis and supervised learning. IEEE Transactions on Intelligent Transportation Systems 14(2), 871–882.

Liu, R., Wang, Y., Zhou, H., Qian, Z., 2019b. Short-term passenger flow prediction based on wavelet transform and kernel extreme learning machine. IEEE Access 7, 158025–158034.

Liu, Y., Liu, Z., Jia, R., 2019a. DeepPF: a deep learning based architecture for metro passenger flow prediction. Transportation Research Part C 101, 18–34.

Lu, W., Ma, C., Li, P., 2020. Research on sample selection of urban rail transit passenger flow forecasting based on SCBP algorithm. IEEE Access 8, 89425–89438.

McFadden, J., Yang, W.T., Durrans, S.R., 2001. Application of artificial neural networks to predict speeds on two-lane rural highways. Transportation Research Record: Journal of the Transportation Research Board 1751, 9–17.

Melo, P.C., Sobreira, N., Goulart, P., 2019. Estimating the long-run metro demand elasticities for Lisbon: A time-varying approach. Transportation Research Part A 126, 360–376.

Milenković, M., Švadlenka, L., Melichar, V., Bojović, N., Avramović, Z., 2018. SARIMA modelling approach for railway passenger flow forecasting. Transport 33(5), 1113–1120.

Ni, M., He, Q., Gao, J., 2017. Forecasting the subway passenger flow under event occurrences with social media. IEEE Transactions on Intelligent Transportation Systems 18(6), 1623–1632.

Nosal, T., Miranda-Moreno, L.F., 2014. The effect of weather on the use of North American bicycle facilities: a multi-city analysis using automatic counts. Transportation Research Part A 66, 213–225.

Park, B., Messer, C., Urbanik II, T., 1998. Short-term freeway traffic volume forecasting using radial basis function neural network. Transportation Research Record: Journal of the Transportation Research Board 1651, 39–47.

Park, D., Rilett, L.R., Han, G., 1999. Spectral basis neural networks for real-time travel time forecasting. Journal of Transportation Engineering 125, 515–523.

Pel, A.J., Bel, N.H., Pieters, M., 2014. Including passengers' response to crowding in the Dutch national train passenger assignment model. Transportation Research Part A 66, 111–126.

Pereira, F.C., Rodrigues, F., Ben-Akiva, M., 2015. Using data from the web to predict public transport arrivals under special events scenarios. Journal of Intelligent Transportation Systems 19(3), 273–288.

Sheng, M., Sharp, B., 2019. Aggregate road passenger travel demand in New Zealand: A seemingly unrelated regression approach. Transportation Research Part A 124, 55–68.

Smith, B.L., Williams, B.M., Oswald, R.K., 2002. Comparison of parametric and nonparametric models for traffic flow forecasting. Transportation Research Part C 10(4), 303–321.

Stathopoulos, A., Karlaftis, M.G., 2003. A multivariate state space approach for urban traffic flow modeling and prediction. Transportation Research Part C 11(2), 121–135.

Sun, Y., Leng, B., Guan, W., 2015. A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system. Neurocomputing 166, 109–121.

Sun, Y., Zhang, G., Yin, H., 2014 Passenger flow prediction of subway transfer stations based on nonparametric regression model. Discrete Dynamics in Nature and Society 2014, Article ID 397154, http://dx.doi.org/10.1155/2014/397154.

Tan, M.-C., Wong, S.C., Xu, J.-M., Guan, Z.-R., Zhang, P., 2009. An aggregation approach to short-term traffic flow prediction. IEEE Transactions on Intelligent Transportation Systems 10(1), 60–69.

Tang, Q., Yang, M., Yang, Y., 2019. ST-LSTM: A deep learning approach combined spatio-temporal features for short-term forecast in rail transit. Journal of Advanced Transportation 2019, Article ID 8392592, https://doi.org/10.1155/2019/8392592.

Tsai, T.H., Lee, C.K., Wei, C.H., 2009. Neural network based temporal feature models for short-term railway passenger demand forecasting. Expert Systems with Applications 36 (2), 3728–3736.

van Lint, J., Hoogendoorn, S., van Zuylen, H., 2002. Freeway travel time prediction with state-space neural networks: modeling state-space dynamics with recurrent neural networks. Transportation Research Record: Journal of the Transportation Research Board 1811, 30–39.

Vlahogianni, E.I., Golias, J.C., Karlaftis, M.G., 2004. Short-term traffic forecasting: overview of objectives and methods. Transport Reviews 24(5), 533–557.

Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2014. Short-term traffic forecasting: where we are and where we're going. Transportation Research Part C 43, 3–19.

Wang, J., Kong, X., Zhao, W., Tolba, A., Al-Makhadmeh, Z., Xia, F., 2018. STLoyal: A spatio-temporal loyalty-based model for subway passenger flow prediction, IEEE Access 6, 47461–47471

Wang, L.-X., 1997. A Course in Fuzzy Systems and Control, Prentice-Hall, Englewood Cliffs, NJ.

Wang, Y., Papageorgiou, M., Messmer, A., 2007. Real-time freeway traffic state estimation based on extended Kalman filter: a case study. Transportation Science 41(2), 167–181.

Wei, Y., Chen, M.C., 2012. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. Transportation Research Part C 21(1), 148–162.

Williams, B.N., Hoel, L.A., 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. Journal of Transportation Engineering 129, 664–672.

Wu, C.H., Ho, J.M., Lee, D.T., 2004. Travel-time prediction with support vector regression. IEEE Transactions on Intelligent Transportation Systems 5, 276–281.

Wu, J., Liao, H., 2020. Weather, travel mode choice, and impacts on subway ridership in Beijing. Transportation Research Part A 135, 111–126.

Xiong, Z., Zheng, J., Song, D., Zhong, S., Huang, Q., 2019. Passenger flow prediction of urban rail transit based on deep learning methods. Smart Cities 2019(2), 371–387.

Yang, D., Chen, K., Yang, M., Zhao, X., 2019. Urban rail transit passenger flow forecast based on LSTM with enhanced long-term features. IET Intelligent Transport Systems 13(10), 1475-1482.

Yu, H.-T., Jiang, C.-J., Xiao, R.-D., Liu, H.-O., Lv, W., 2019. Passenger flow prediction for new line using region dividing and fuzzy boundary processing, IEEE Transactions on Fuzzy Systems 27(5), 994–1007.

Zadeh, L.A., 1965. Fuzzy sets. Information and Control 8(3), 338–353.

Zadeh, L.A., 1972. A fuzzy-set-theoretic interpretation of linguistic hedges. Journal of Cybernetics 2, 4–34.

Zadeh, L.A., 1973. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on Systems, Man, and Cybernetics 3, 28–44.

Zeng, D., Xu, J., Gu, J., Liu, L., Xu, G., 2008. Short term traffic flow prediction using hybrid ARIMA and ANN models. Proceedings of 2008 Workshop on Power Electronics and Intelligent Transportation System, pp. 621–625.

Zhang, A., Wan, Y., Yang, H., 2019b. Impacts of high-speed rail on airlines, airports and regional economies: A survey of recent research. Transport Policy 81, A1–A19.

Zhang, D., Wang, X., 2014. Transit ridership estimation with network kriging: a case study of second avenue subway, NYC. Journal of Transport Geography 41, 107–115.

Zhang, F., Graham, D.J., Wong, M.S.C., 2018b. Quantifying the substitutability and complementarity between high-speed rail and air transport. Transportation Research Part A 118, 191–215.

Zhang, H.M., 2000. Recursive prediction of traffic conditions with neural network models. Journal of Transportation Engineering 126(6), 472–481.

Zhang, J., Chen, F., Shen, Q.,2019a. Cluster-based LSTM network for short-term passenger flow forecasting in urban rail transit. IEEE Access 7, 147653–147671.

Zhang, Q., Mahfouf, M., 2007. A new reduced space searching algorithm (RSSA) and its application in optimal design of alloy steels. Proceedings of the 2007 IEEE Congress on Evolutionary Computation, pp. 1815–1822.

Zhang, Q., Mahfouf, M., 2008. Mamdani-type fuzzy modelling via hierarchical clustering and multi-objective particle swarm optimisation (FM-HCPSO). International Journal of Computational Intelligence Research 4(4), 314–328.

Zhang, Q., Mahfouf, M., 2010. A nature-inspired multi-objective optimisation strategy based on a new reduced space searching algorithm for the design of alloy steels. Engineering Applications of Artificial Intelligence 23(5), 660–675.

Zhang, Y., Ye, Z., 2008. Short-term traffic flow forecasting using fuzzy logic system methods. Journal of Intelligent Transportation Systems 12(3), 102–112.

Zhang, Z., He, Q., Gao, J., Ni, M., 2018a. A deep learning approach for detecting traffic accidents from social media data. Transportation Research Part C 86, 580–596.

Zhang, Z., Wang, C., Gao, Y., Chen, Y., Chen, J., 2020. Passenger flow forecast of rail station based on multi-source data and long short term memory network. IEEE Access 8, 28475–28483.

Zhao, S., Mi, X., 2019. A novel hybrid model for short-term high-speed railway passenger demand forecasting. IEEE Access 7, 175681–175692.

Zhao, Y., Ma, Z., Yang, Y., Jiang, W., Jiang, X., 2020. Short-term passenger flow prediction with decomposition in urban railway systems. IEEE Access 8, 107876–107886.

Zhu, K., Xun, P., Li, W., Li, Z., Zhou, R., 2019. Prediction of passenger flow in urban rail transit based on big data analysis and deep learning. IEEE Access 7, 142272–142279.