

Liyanage, Sohani; Abduljabbar, Rusul; Dia, Hussein; Tsai, Pei-Wei

Article

AI-based neural network models for bus passenger demand forecasting using smart card data

Journal of Urban Management

Provided in Cooperation with:

Chinese Association of Urban Management (CAUM), Taipei

Suggested Citation: Liyanage, Sohani; Abduljabbar, Rusul; Dia, Hussein; Tsai, Pei-Wei (2022) : AI-based neural network models for bus passenger demand forecasting using smart card data, Journal of Urban Management, ISSN 2226-5856, Elsevier, Amsterdam, Vol. 11, Iss. 3, pp. 365-380, <https://doi.org/10.1016/j.jum.2022.05.002>

This Version is available at:

<https://hdl.handle.net/10419/271473>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>



Research Article

AI-based neural network models for bus passenger demand forecasting using smart card data

Sohani Liyanage^{a,*}, Rusul Abduljabbar^a, Hussein Dia^a, Pei-Wei Tsai^b^a Department of Civil and Construction Engineering, Swinburne University of Technology, Australia^b Department of Computer Science and Software Engineering, Swinburne University of Technology, Australia

ARTICLE INFO

Keywords:

Artificial intelligence
Short-term prediction
Neural networks
Bus demand prediction
Deep learning
On-demand public transport

ABSTRACT

Accurate short-term forecasting of public transport demand is essential for the operation of on-demand public transport. Knowing where and when future demands for travel are expected allows operators to adjust timetables quickly, which helps improve service quality and reliability and attract more passengers to public transport. This study addresses this need by developing AI-based deep learning models for prediction of bus passenger demands based on actual patronage data obtained from the smart-card ticketing system in Melbourne. The models, which consider the temporal characteristics of travel demand for some of the heaviest bus routes in Melbourne, were developed using real-world data from 18 bus routes and 1,781 bus stops. LSTM and BiLSTM deep learning models were evaluated and compared with five conventional deep learning models using the same data set. A desktop comparison was also undertaken against a number of established demand forecasting models that have been reported in the literature over the past decade. The comparative evaluation results showed that BiLSTM models outperformed other models tested and was able to predict passenger demands with over 90% accuracy.

1. Introduction

A well-developed urban public transport system, especially bus transport, can reduce congestion and emissions and decrease the use of private vehicles (Li, Cao, et al., 2020). On-demand public transport, in particular, is seen to have the potential to improve operations further and enhance customer satisfaction. However, this type of service requires that short-term forecasts of future demands for bus services are known in advance (Liang et al., 2019; Liyanage & Dia, 2020; Liyanage et al., 2019; Smith et al., 2002; Zhou et al., 2013). Accurate prediction of future demands also helps operators to pre-allocate constrained resources such as vehicles and drivers to meet passenger demands and provide quality and reliable services with minimum waiting times. It also allows operators to optimize bus fleet management to minimise operational costs (Ma et al., 2014; Tirachini et al., 2013). Demand prediction is an integral part of business and commerce operations and helps decision makers to reduce the uncertainties of future operations. In public transport operations, the passenger service business models are highly dependent on accurate estimation of future passenger demands. Starting from route design and network planning, through to scheduling of vehicles with optimised seating capacity to meet operators' and users' objectives, and pricing each passenger vehicle on a network route, service operations in every planning horizon is dependent in one way or another on accurate estimation of future demands (Banerjee et al., 2020; Lu et al., 2021).

The focus of this paper is directly aligned with current advancements in digitalisation of urban mobility planning tasks. This is

* Corresponding author. Swinburne University of Technology, Melbourne, Australia.

E-mail address: sliyanage@swin.edu.au (S. Liyanage).

manifested through the paper's focus on development and evaluation of advanced passenger demand forecasting models which are important pillars in the planning and delivery of efficient and customer-focused on-demand public transport systems. First, the paper presents developments in app-based on-demand public transport services that are supported by technologies and advancements in IoT, Big Data and Real-Time Analytics. The paper then outlines how the success of these emerging modes of on-demand public transport, currently being run in trials in a number of cities around the world (including Sydney), relies heavily on accurate estimations of travel demand and passenger numbers over short durations up to 1 h. The paper then presents development and evaluation of advanced methodologies for passenger demand forecasting based on AI and deep learning theories that can be used for estimating future passenger demands. Another novel aspect of this work directly aligned to digitalisation of urban infrastructure is the use of smart card based field passenger demand data for model development. Unlike a large number of studies on this topic that used simulated data, this paper uses real-world field data obtained from IoT-enabled MyKi fare collection smart cards used for public transport in Melbourne.

Long-term public transport demand forecasting methodologies, including the well-established four-step models, elasticity and economic models, long-range demand models, mainly focus on planning issues and are not suited for operational forecasts (Balcombe, 2004; Ma et al., 2014). The research on short-term passenger demand forecasts for operational purposes is far more limited. Unlike the planning models, short-term passenger demand prediction models are used to account for short-term dynamic changes in demand due to weather conditions, traffic congestion, and special events. These models have gained widespread attention from transport planners and researchers in recent years due to their important operational role (Dia et al., 2001; Noekel, 2016; Zhai et al., 2018), making them a prerequisite for proactive operations and management of bus transport services (Ceder et al., 2013; Xue et al., 2015).

Demand prediction methodologies are based on understanding the temporal and spatial relationship between different variables in historical data. The primary objective is to obtain accurate and realistic forecasting of future demands (Zhai et al., 2020). Parametric and non-parametric methods are generally the two main solution techniques for short-term demand prediction techniques (Wei & Chen, 2012; Wu, Jiang, et al., 2020). In parametric approaches (also known as linear models or statistic approaches), the Box-Jenkins methodology is the most widely used model (Box et al., 2015). This model applies ARIMA or ARMA, AR, decision tree models to identify trends (Anvari et al., 2016; Cyril et al., 2018; Gan et al., 2014; Gong et al., 2014; Milenković et al., 2018; Wu, Xia, & Jin, 2020). However, these models are limited in scope because they are developed based on linear assumptions among time-lagged variables (Bai et al., 2017). Examples where these linear approaches were applied include two studies by (Ma et al., 2014) and (Xue et al., 2015) where the authors proposed Interactive Multiple Models that combine different algorithms to forecast passenger demands during different times of the day. In the first study (Ma et al., 2014), applied AR, SARIMA and ARIMA for weekly, daily, and hourly time-series analyses, respectively. In the second study (Xue et al., 2015), applied ARMA SARIMA and ARIMA models for weekly, daily and hourly time-series analyses, respectively. However, the linear models depend on high-quality data comprising accurate and non-fluctuating patterns to develop a time-series sequence. However, real-world passenger demand data are random and unstable, which renders linear approaches incapable of describing the variations in passenger flows. Hence, other methodologies have been presented in the literature to track such non-linear characteristics.

The second category, non-linear or non-parametric approaches, constructs non-linear relationships between input and output variables (Wu, Li, et al., 2020). There are also hybrid approaches that combine multiple algorithms strategically. Non-parametric models include support vector machine (SVM) (Chen et al., 2015; Jiang et al., 2014; Yang & Liu, 2016), least-squares SVM (Guo et al., 2013; Sun et al., 2015; Zhang et al., 2011), Fuzzy neural networks (Buckley & Hayashi, 1994; Dou et al., 2013; Tsai et al., 2009), Bayesian networks (Roos et al., 2017; Sun et al., 2006), grey models (Hai-lan & En-chong, 2012; Wang & Zhang, 2012; Yang et al., 2013; Zhang et al., 2017) and neural networks (Chen et al., 2012; Pekel & Soner Kara, 2017; Zhang et al., 2013). Among these, Artificial Intelligence (AI) approaches, based on deep learning neural network methodologies, have been identified as most promising and practical for complex time series forecasts (Lee et al., 2006) and have been shown to provide improved predictive capability (Vlahogianni & Karlaftis, 2011).

Other recent literature on passenger demand forecasting also examined the relationship between Internet of Things sensing infrastructures, intelligent vehicular networks, and big data-driven algorithmic decision-making for provision of accurate short-term forecasting of public transport demand (Aldridge & Stehel, 2021; Blake & Michalikova, 2021; Lăzăroi & Harrison, 2021; Nica, 2021). Recent literature on passenger demand forecasting also examined the relationship between deep learning-based sensing technologies, predictive control algorithms, and big geospatial data analytics (Adams et al., 2021; Pelau et al., 2021; Rowland & Porter, 2021; Wallace & Lăzăroi, 2021). Along with the above, AI-based neural networks have attracted research interest in a large number of transport application areas (Dia, 2001; Dia et al., 1996; Dia & Panwai, 2011; Dia & Rose, 1995; Thomas & Dia, 2006; Thomas et al., 2001). Recently, deep learning neural network innovations have also attracted research interest in the context of demand prediction. For example, in a study by (Liu et al., 2019, 2020), the authors predicted short-term metro passenger flows using end-to-end deep learning architectures using multiple sources of data. A “multi-pattern deep fusion (MPDF)” approach was proposed by (Bai et al., 2017), which was constructed through fusing “Deep Belief Networks (DBNs)” to represent passenger flow, where the DBNs were proven to be successful in extracting complex features (Li, Cabrera, et al., 2016). An “End-to-end Multi-task Learning Temporal CNN” was also proposed by (Zhang et al., 2019) to forecast short-term passenger demand at a multi-zone levels. A hybrid estimation model was introduced based on CNN and spatial-temporal contexts (Liu & Chen, 2017; Liu et al., 2017). In their model, the CNN detected passengers, while the spatial-temporal model accurately predicted passenger volumes. Another study captured spatial-temporal correlations of demands using a “Fusion Convolutional-LSTM deep learning” approach (Ke et al., 2017). Also (Han et al., 2019), proposed an optimised hybrid LSTM model for accurate bus passenger flow prediction, which was shown to have excellent performance accuracies. LSTM models have also been used in spatial-temporal aspects to forecast dynamic origin-destination matrices in a subway network in France (Toqué et al., 2016). In another study, a spatial-temporal model was redesigned for rail transit passenger demand estimation based on multiple data sources such as smart card, weather and mobile phone data (Li, Cao, et al., 2020). Similarly, origin-destination matrices prediction was tested in the context of ride-hailing or taxi passenger demand utilising various non-parametric models including LSTM approaches

Table 1
Description of the fields of the dataset.

Field	Description
Smart card touch-on	Transaction time of touch-on, including date, in POSIX format (AEST)
Smart card touch-off	Transaction time of touch-off, including date, in POSIX format (AEST)
Origin bus stop	Stop ID of boarding
Route number	Route ID of boarding
Route direction	Route direction of boarding
Destination bus stop	Stop ID of alighting
Paired type	Not every touch-on has a corresponding touch-off. Some touch-offs are estimated through a model that uses existing 'paired' trips to infer the likely 'synthetic' touch-off. Trips that are 'paired' are more reliable than 'synthetic.'
Boost factor	Boost factor used to scale up the transaction to approximate actual patronage

(Chen et al., 2020; He et al., 2022; Ke et al., 2021; Wang et al., 2019; Zhang, Che, et al., 2021; Zhang, Che, et al., 2021).

Based on the results reported in previous studies, most parametric and non-parametric methods faced major challenges in producing the high accuracies (90% and above) required for reliable real-world operations. This study bridges these gaps by developing deep learning passenger bus transport demand forecasting models with required accuracy in the short-term context (15-min, 30-min and 60-min time horizons). As shown in the results section, the accuracy of our models exceeded 90% for all tested scenarios. Generally, neural network models' performance is challenged by difficulties in feature learning and complexities in pattern recognition (Panwai & Dia, 2006). Researchers have used data pre-processing for pattern recognition (Bai et al., 2017; Panwai & Dia, 2005; Rose & Dia, 1995) and intensive networks for feature learning (Hinton et al., 2006). In our study, which utilised the neural network approach to predict passenger demand on bus routes, time-series decomposition is used to pre-process the data and to understand various patterns. Deep learning neural networks (i.e. networks with more than one hidden layer) were then used for intensive feature learning (Bengio, 2009; Bengio et al., 2007; Li, Bai, & Zeng, 2016).

From a practical perspective, and in addition to the academic contributions of this paper, it is important to note that passenger demand forecasting is an integral part of real-world public transport operations. Decision makers need to have accurate estimates of what the demand for travel will be particularly for short-term horizons. Knowing these patterns of demands helps operators to optimize operations, reduce costs and enhance customer satisfaction. This study is therefore directly related to addressing this industry challenge by using state-of-the-art AI-based deep learning forecasting models that improve passenger bus transport demand forecasting for short-term horizons up to 60 min into the future. The model developments and enhancements discussed in this study, which represent meaningful advancements on today's best practice algorithms and solutions, have both a technical and academic significance in addition to important practical significance to the industry.

This study aims to improve passenger bus transport demand forecasting in the short-term context (15-min, 30-min and 60-min time horizons). The key research gap addressed in this paper is the lack of accurate public transport passenger demand forecasting models that can be used reliably for real-world applications. While previous literature relied on development of statistical-based forecasting models using simulation data, this paper presents state-of-the-art deep learning models that were developed using real-world IoT-enabled smart card fare collection data reflecting real-world conditions. This contributed to development of highly reliable and accurate models compared to methodologies described in existing literature.

The remainder of this paper includes the following sections. The second section presents the data sources used for model development, specification of the study area and data requirements, methodological framework, time-series pattern development and necessary data preparations to facilitate demand prediction. Section 3 presents the LSTM model development for selected bus routes and reports on a comparative analysis of a number of neural network models. The fourth section presents an analysis of results and compares the performance of the tested models. Finally, the fifth section presents conclusions and insights for further research work.

2. Data source and methodology

This study relied on passenger demand data from Melbourne's MyKi smart card ticketing system, which is a reloadable contactless smart card launched in 2008. The system became fully operational for electronic payment of public transport fares in 2012. The Myki card system is convenient for passengers with easy access to fare payment devices that are installed in public service vehicles. Passengers can immediately pay for their trips by touching on and off from their start and end locations. For this study, the authors' accessed one-month data, May 2018, from the Department of Transport in Victoria.

2.1. Data format and fields

The MyKi data is collected from recorded passenger touch-on and touch-off events and used to obtain passenger boarding and alighting counts for bus service routes. This data and information about Metropolitan Melbourne bus routes in the study area were also obtained. The dataset received for this study contained the fields presented in Table 1. Approximately 10–14% of bus transactions occurred on transport vehicles where MyKi was operating in a "headless mode", meaning that the service operator was known, but the

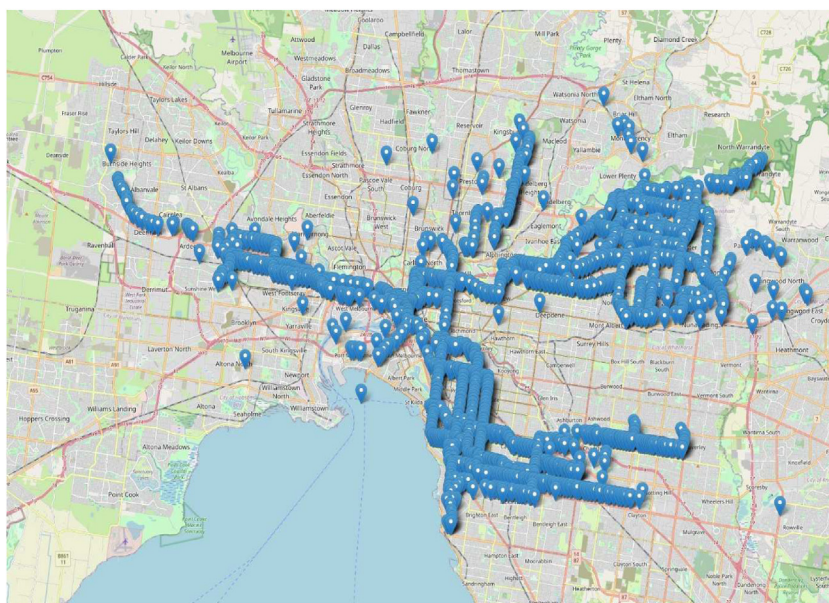


Fig. 1. Distribution of bus stop locations covering 18 routes in the study area.

Table 2

Details of the bus routes operating in the study area.

Route ID	Service Direction	Number of bus stops	Total trips
200	City (Queens Street) via Kew Junction	58	44,237
	Bulleen via Kew Junction	60	
207	City (Queens Street) via Kew Junction	58	50,445
	Doncaster SC via Kew Junction	61	
216	City (Queens Street) Brighton Beach	87	18,932
	Sunshine Station	86	
219	Gardenvale via City	87	19,496
	Sunshine South via City	92	
220	Gardenvale via City	88	42,868
	Sunshine via City	88	
246	Clifton Hill via St Kilda	45	65,541
	Elsternwick via St Kilda	44	
302	City (Lonsdale St) via Belmore Rd > Eastern Freeway	51	42,917
	Box Hill via Eastern Freeway > Belmore Road	47	
304	City (Lonsdale St) via Belmore Rd > Eastern Freeway	47	26,001
	Doncaster SC via Eastern Freeway > Belmore Road	45	
305	City (King/Lonsdale Streets) via Eastern Freeway	50	24,132
	The Pines via Eastern Freeway	48	
309	Donvale via Eastern Freeway > Reynolds Road	53	10,395
	City (Queens street) via Reynolds Road > Eastern Freeway	54	
350	La Trobe University via Eastern Freeway	41	12,679
	City (Queen Street) via Eastern Freeway	45	
605	Flagstaff Station via Kooyong Road	53	13,337
	Gardenvale via Kooyong Road	49	
623	St Kilda via Chadstone SC	72	28,674
	Glen Waverley via Chadstone SC	73	
625	Elsternwick via Oakleigh > Ormond	50	9,652
	Chadstone SC via Ormond > Oakleigh	51	
630	Elwood via Ormond	54	57,292
	Monash University via Ormond	53	
905	City (King/Lonsdale Streets) via Templestowe > Eastern Freeway	47	56,409
	The Pines via Eastern Freeway > Templestowe	43	
906	City (Kind/Lonsdale Streets) via The Pines SC	51	68,906
	Warrandyte via The Pines SC	48	
907	City (Kind/Lonsdale Streets) via Doncaster Road > Eastern Freeway	47	101,776
	Mitcham via Eastern Freeway > Doncaster Road	46	

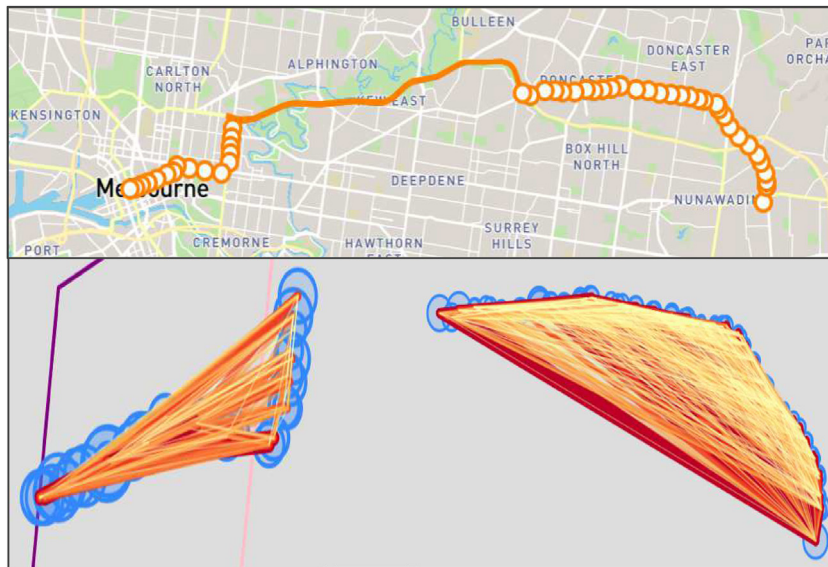


Fig. 2. (a) Map of movement of route 907 in both service directions (b) Bus trip distribution among bus stops along route 907.

route was unknown. It is also estimated that approximately 60% of passengers generally touch their cards when using buses, but a lower percentage touches off. As a result, while smart card readers may not record every traveller's information, the data can still be used to reflect and understand passenger travel demand. The raw passenger origin-destination data was supplied to the research team in an anonymised format. School buses were excluded from this analysis.

To account for passengers who may not touch on or off, missing passenger patronage was estimated by the Department of Transport using surveys to supplement smart card data. Passengers may not touch on for various reasons, such as faulty smart cards, faulty smart card readers or fare evasion. The survey data was used to arrive at a weighting (or boost factor) associated with each transaction to approximate actual bus patronage. The dataset was thoroughly checked, and all invalid trips were removed.

The smart card data provided a good understanding of existing bus passenger demand profiles and patterns, which in turn was sufficient for producing accurate and reliable prediction results. The data set included information for the entire bus route (18-service routes with both service directions for each route) and 1,781 touch-on bus stops. The specific routes included routes 200, 207, 246, 302, 304, 305, 309, 350, 905, 906, 907, 605, 623, 630, 220, 216, 219, and 625 in both service directions. The operating timetables for existing bus services in Metropolitan Melbourne (during which the MyKi data was obtained) were extracted from Public Transport Victoria travel information portal. They included separate schedules for weekdays, Saturdays and Sundays (PTV, 2019).

2.2. Study area and data description

The city of Melbourne in Australia was selected as the case study for this research. Melbourne has one of Australia's largest urban population, comprising more than 5 million people (World population review, 2021). Each blue colour place marker in Fig. 1 represents an existing bus stop for the 18 routes covered in the available data across the Southeast region of Melbourne.

The study area included a number of inner suburbs in the southeast of Melbourne, which covered significant bottleneck arterials that are among the most congested corridors in Melbourne. The case study area is also well-served by different types of public transport, including fixed-service buses equipped with MyKi smart-card readers. Among 715,375 trips, 693,689 trips were useable (96.97%) for this analysis. The study area covered 18-service routes and 1,781 bus stops in the Metropolitan Melbourne region, summarised in Table 2.

In the Metropolitan Melbourne area, bus services in both service directions move intensively between 5:00 a.m. to 1:00 a.m. The traffic maps for each of the 18 bus routes were developed to visualise where they run within the study area and to also understand the total passenger trip distributions between bus stops. For example, Route 907 (the most attractive public bus transport movement among the 18 service routes) is selected to present a zoomed view of the details (Fig. 2). Route 907 is shown in Fig. 2a, where each stop is marked with a circle whose size is proportional to number of trips originating from and ending at that stop. The width of the line linking each origin to destination represents the number of passenger journeys between the bus stops. There were 78 service vehicles on the 907-route and 77 vehicles on the 907b route for an average weekday. Route 907 is 36 km in length and traverses mainly arterials and the Eastern Freeway. Residential communities, schools and workplaces are located along this bus route, accommodating many work and education trips on weekdays.

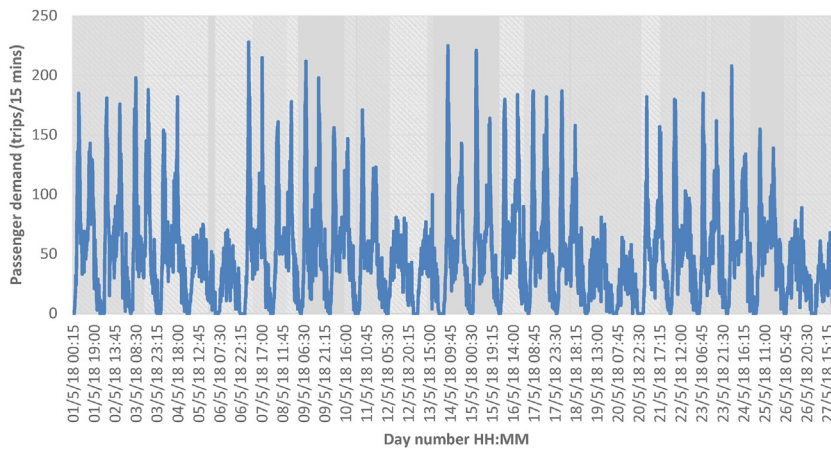


Fig. 3. Passenger demand for Route 907 in May 2018.

Table 3
Heat map of passenger counts for route 907.

Date/Time	1:00	6:00	7:00	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	0:00
Monday	0	19	113	390	436	191	156	108	131	140	167	278	276	404	240	125	98	69	37	23
Tuesday	2	33	162	514	613	257	206	190	225	223	249	350	364	465	299	168	113	79	58	23
Wednesday	7	27	160	588	528	191	194	180	218	242	243	347	387	566	317	136	133	132	83	52
Thursday	8	26	154	538	537	201	222	194	215	225	254	422	368	520	285	132	136	118	55	24
Friday	11	23	140	468	500	222	224	177	200	225	225	376	360	463	276	146	150	157	86	79
Saturday	26	0	36	85	161	165	193	206	216	189	183	192	198	224	158	107	75	69	75	36
Sunday	5	0	2	49	113	139	189	204	160	139	137	143	161	172	88	76	83	21	0	0

2.3. Data structure and methodology

The methodology adopted in this study is based on an empirical data-driven framework for predicting bus passenger demands using neural networks. The first step includes dividing the full data set into three-time series, namely 15, 30 and 60-min. The second step was to develop corresponding neural network models to capture the series features and generate reliable and accurate prediction outputs. The demand represents the number of passengers in both directions.

The historical data are aggregated in 15 min, as 15 min is the most extended departure interval experienced for all 18 routes during off-peak periods. There were 80-time intervals between 05:00 to 24:00 + 1 (i.e. 1:00 the following day). To better understand the characteristics of passenger demand, the historical data in May 2018 was plotted for all 18 routes.

Fig. 3 provides an example for Route 907. There are no public or school holidays in this selected month of 2018. The line diagram in Fig. 2 shows passenger demand fluctuation represents cyclical variations that follow a similar pattern almost every week for all weekdays and weekends. It is interesting to note that significant gaps in passenger demand are observed from peak to off-peak periods on weekdays, and from weekdays to weekends due to reduced work and study trips. Furthermore, pronounced asymmetries of demand of bus passengers were identified between morning and evening peaks. Among these, the morning peak is characterised by sharp and pronounced spikes compared to the evening peak because work and study journeys usually coincide in the morning. In contrast, the afternoon peak is flatter and more spread due to earlier return-to-home study trips. Not surprisingly, demand fluctuates less during weekends.

Table 3 shows a heat map of the number of passengers per hour averaged over seven days of the week for route 907. The colours indicate where the value sits in different ranges: Red represents the lowest passenger count during each hour of an average day, while green depicts the highest passenger count.

Three prediction models are developed for 15, 30, and 60 min prediction horizons. A description of the adopted for this work is presented next.

A vector of a single row of elements for the time series for day n is represented as $d^n = \{d^n(t) | t = 1, 2, \dots, N_{Tin}\}$. In this, $d^n(t)$ is the bus passenger demand at a time interval $(t - 1, t]$ for day n, $t \in [1, 80]$. Here, t represents an integer and N_{Tin} the number of time intervals in a day (80-time intervals for this study).

Then in the source dataset, a time series is defined as $d_s = [d^1, d^2, \dots, d^{N_{day}}]$. N_{day} represents the full days covered in the dataset, ranging from 1st to 28th May for this study. Relevant pattern time series of three different time series was constructed, which are 15-

Table 4
Demand datasets used for each prediction horizon.

Time horizon	Total dataset	Training dataset	Test and validation dataset
15-min	27,823	16,694	11,129
30-min	27,360	16,416	10,944
60-min	26,561	15,937	10,624

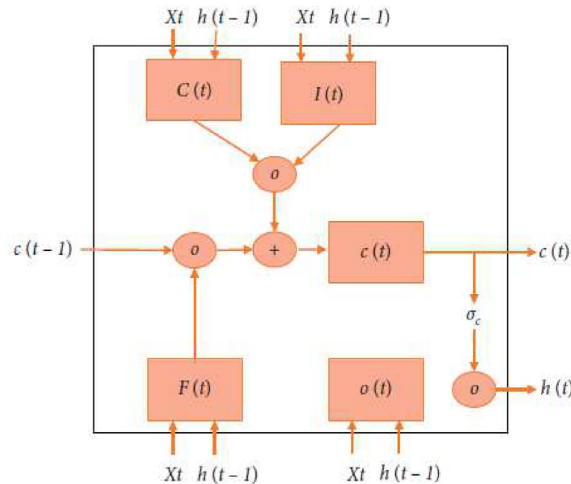


Fig. 4. LSTM model architecture.

min, 30-min, and hourly.

1. For 15min: $TS_{15m}^n(t)$ consists of bus demand data $d^n(t)$ in n_{15m} time intervals.

$$TS_{15m}^n(t) = \{d^n(t-1), d^n(t-2), \dots, d^n(t-n_{15m})\}$$

2. For 30min: $TS_{30m}^n(t)$ consists of bus demand data $d^n(t)$ in n_{30m} time intervals.

$$TS_{30m}^n(t) = \{d^n(t-1), d^n(t-2), \dots, d^n(t-n_{30m})\}$$

3. For 60min: $TS_h^n(t)$ consists of bus demand data $d^n(t)$ in n_h time intervals.

$$TS_h^n(t) = \{d^n(t-1), d^n(t-2), \dots, d^n(t-n_h)\}$$

2.4. Data preparation for neural network modelling

Passenger demand data collected from the smart card data system were aggregated for different durations such as 15 min, 30 min, and 1 h. It is important to note that temporal data aggregation can alter the underlying characteristics and features of the source field data (Vlahogianni & Karlaftis, 2011). For this study, 15-min data aggregation was selected because it was the maximum interval for bus departures during off-peak periods. The 15 min aggregated passenger demand was organised by day and by week from Monday to Sunday over the entire month. Among the 80-time intervals, from 05:00 to 24:00 + 1 (i.e. 1:00 the following day), erroneous samples and time intervals with significantly less passenger demand were excluded. Therefore, only time intervals from morning 07:00 to evening 21:30 within a day were considered. Model development included dividing the available data into “training and testing” data sets. The training set was approximately 70% of the whole dataset and was used to develop the model to calibrate the model parameters. The testing data set comprised about 30% of the total dataset used for the model validation and to test the demand prediction accuracy of the developed models. The data sets used in every three models are presented in Table 4.

Neural network models work poorly with data sets that have high variability. For example, in our dataset, passenger demand changes between 0 to around 250 passengers' trips per 15-min. Such heterogeneity of data can complicate the learning process (Faridai et al., 2021). The heterogeneity of data was addressed using data smoothing, which was used to detect trends in noisy data in cases where the shape of the movement is unknown. Exponential smoothing was used with a 0.5 damping factor.

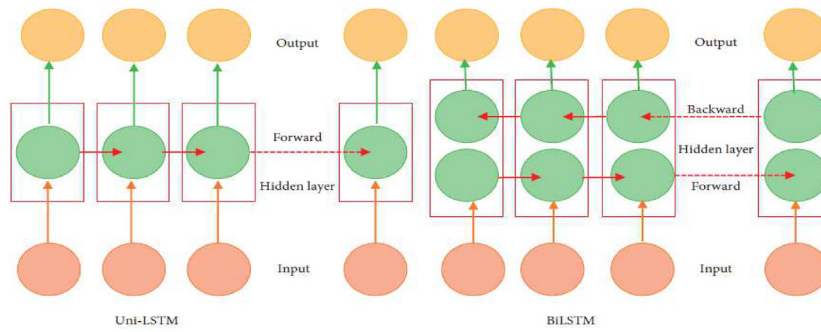


Fig. 5. Uni-LSTM and BiLSTM architectures.

Table 5
LSTM model parameters.

Parameters	Settings
Gradient Decay Factor	0.9
Initial Learning Rate	5*10 ⁽⁻³⁾
Minimum Batch Size	128
Maximum Epochs	300
Training Optimizer	Adaptive moment estimation optimizer
Dropping Learning Rate During Training	Piecewise
Learning Rate Drop Period	125
The factor for Learning Rate Dropping	0.2

3. Artificial neural network model development

Recurrent Neural Networks (RNNs) architectures were in the past found to provide good prediction accuracy. However, with the inability to use information from a distant past, RNNs did not perform well for long-term memory (Abduljabbar, Dia, Tsai, et al., 2021). The LSTM models address this problem as they are considered extensions of the RNNs (Hochreiter & Schmidhuber, 1997). LSTM models have the capability of learning patterns with long dependencies compared to traditional RNNs that are not known for performing well for longer horizon patterns. Hence, LSTM models are generally identified to outperform RNNs in time series data forecasting (Yeon et al., 2019). Uni-directional and bi-directional LSTM models were developed in this study. Bi-directional LSTM (BiLSTM) is an extension to the unidirectional LSTM that includes additional data training where the model trains in forward and backward directions (Abduljabbar, Dia, & Tsai, 2021). The model architectures are presented below. LSTM models have different structures than RNN models, as presented in Fig. 4 and Fig. 5. The following formulae calculate the predicted values.

$$\text{Input gate } (I_t) = \sigma_g(W_i X_t + R_i h_{t-1} + b_i) \tag{1}$$

$$\text{Forget gate } (f_t) = \sigma_g(W_f X_t + R_f h_{t-1} + b_f) \tag{2}$$

$$\text{Cell Candidate } (C_t) = \sigma_g(W_c X_t + R_c h_{t-1} + b_c) \tag{3}$$

$$\text{Output gate } (O_t) = \sigma_g(W_o X_t + R_o h_{t-1} + b_o) \tag{4}$$

Where:

- σ_g - Gate activation function
- $W_{i,f,c,o}$ - Input weight matrices
- $R_{i,f,c,o}$ - Recurrent weight matrices
- X_t - Input
- h_{t-1} - Output at the previous time $t - 1$
- $b_{i,f,c,o}$ - Bias vectors

The forget gate determines the “levels of prior memory that need to be removed from the cell state” (Yeon et al., 2019). In similar manner, the input gate specifies new (cell state) inputs. Then, the cell state C_t and the output H_t at time, t is calculated as follows for the LSTM.

Table 6
Neural network architectures and parameters.

Model Architecture	Model Parameters	Learning Rule	Transfer Function	
Deep Learning Backpropagation (DLBP)	LCoeff (Learning rate)	Delta-Rule	Linear	
	Hidden layer 1–0.3	Norm-Cum-Delta	TanH	
	Hidden layer 2–0.25	Ext DBD	Sigmoid	
	Hidden layer 3–0.2	Quick Prop	DNNA	
	Output – 0.15	Max-Prop	Sine	
	Momentum 0.4	Delta-Bar-Delta		
	Trans. Pt. 10000			
	LCoeff Ratio 0.5			
	F' Offset 0.1			
	Epoch 16			
Modular Neural Network (MNN)	LCoeff (Learning rate)	Delta-Rule	Linear	
	Hidden layer 1–0.3	Norm-Cum-Delta	TanH	
	Output – 0.15	Ext DBD	Sigmoid	
	Momentum 0.4	Quick Prop	DNNA	
	Trans. Pt. 10000	Max-Prop	Sine	
	LCoeff Ratio 0.5	Delta-Bar-Delta		
	F' Offset 0.1			
	Epoch 4			
	Radial Basis Function Network (RBFN)	LCoeff (Learning rate)	Delta-Rule	Linear
		Proto – 0.3	Norm-Cum-Delta	TanH
Hidden layer 2–0.25		Ext DBD	Sigmoid	
Output – 0.15		Quick Prop	DNNA	
Momentum 0.4		Max-Prop	Sine	
Trans. Pt. 10000		Delta-Bar-Delta		
LCoeff Ratio 0.5				
Map Trans. 2000				
Epoch 16				
Summation Function - Euclidean				
General Regression Neural Network (GRNN)	Tau (time constant) – 1000	NA	NA	
	Reset Factor – 0			
	Radius of Influence – 0.25			
	Sigma Scale – 1			
	Sigma Exponent – 0.5			
	Summation Function – Euclidean			
Recurrent Neural Networks (RNNs)	LCoeff (Learning rate)	Delta-Rule	Linear	
	Hidden layer 1–0.3	Norm-Cum-Delta	TanH	
	Hidden layer 2–0.25	Ext DBD	Sigmoid	
	Hidden layer 3–0.2	Quick Prop	DNNA	
	Output – 0.15	Max-Prop	Sine	
	Momentum 0.4	Delta-Bar-Delta		
	Trans. Pt. 10000			
	LCoeff Ratio 0.5			
	F' Offset 0.1			
	Epoch 16			

$$C_t = ft \odot ct \odot 1 + it \odot gt \tag{5}$$

$$H_t = ot \odot \sigma c(ct) \tag{6}$$

Where; \odot represents the Hadamard product. This product denotes the element-wise vector multiplication.

For this work, the uni-directional and bi-directional LSTM models were developed in Matlab (Matlab, 2021). The data for short-term temporal prediction of bus routes were first arranged as two-column values: passenger demand at the time (t) and expected passenger demand (t + n), where n was 15, 30 and 60-min. The training (70% of the whole dataset) and testing (30% of the whole dataset) were standardised to prevent the model from overfitting. The developed LSTM network consisted of four layers, including a sequence input layer with one feature, uni-LSTM and bi-LSTM layers with 300 hidden units, a fully connected layer with one response and a regression layer. This developed model can be used in future research work for short-term demand prediction for any aspect where the dataset shows a time-series pattern.

The model settings are as presented in Table 5. Several trials were undertaken to obtain the most-suited combination of values to reach the highest accuracy. For the state activation function, tanH is used, and the sigmoid function is used for the gate activation function. The uniLSTM and biLSTM model development experiments were developed with Deep Learning Toolbox in Matlab R2019b.

3.1. Model evaluation

In order to evaluate the robustness of the LSTM prediction model, the accuracy of the prediction is compared with five other neural

Table 7
Short-term demand prediction accuracies for bus routes.

Bus Route		15-minutes			30-minutes			Hourly			
Prediction horizon		MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	
Temporal prediction accuracies for bus routes	BiLSTM	Error	2.99%	0.78%	2.57%	5.23%	1.60%	3.81%	12.39%	5.62%	8.19%
		Accuracy	97.01%	99.22%	97.43%	94.77%	98.40%	96.19%	87.61%	94.38%	91.81%
	LSTM	Error	22.51%	14.32%	20.11%	30.02%	17.85%	24.98%	35.28%	18.74%	25.71%
		Accuracy	77.49%	85.68%	79.89%	69.98%	82.15%	75.02%	64.72%	81.26%	74.29%
	RNN	Error	25.92%	19.63%	29.88%	34.91%	28.65%	46.34%	48.66%	42.88%	72.45%
		Accuracy	74.08%	80.37%	70.12%	65.09%	71.35%	53.66%	51.34%	57.12%	27.55%
	DLBP	Error	23.79%	19.71%	30.20%	54.21%	32.90%	48.23%	45.06%	42.47%	71.94%
		Accuracy	76.21%	80.29%	69.80%	45.79%	67.10%	51.77%	54.94%	57.53%	28.06%
	MNN	Error	23.79%	19.71%	30.20%	54.21%	32.90%	48.23%	45.06%	42.47%	71.94%
		Accuracy	72.76%	80.04%	69.22%	67.84%	70.48%	51.35%	51.88%	59.33%	31.87%
	RBF	Error	26.73%	23.77%	46.13%	32.90%	32.59%	56.52%	59.53%	46.63%	77.04%
		Accuracy	73.27%	76.23%	53.87%	67.10%	67.41%	43.48%	40.47%	53.37%	22.96%
	GRNN	Error	71.01%	30.50%	49.51%	78.33%	38.04%	58.96%	72.40%	46.31%	75.17%
		Accuracy	28.99%	69.50%	50.49%	21.67%	61.96%	41.04%	27.60%	53.69%	24.83%

network architectures based on the same data set. These included “Deep Learning Back-Propagation (DLBP)”, “Modular Neural Networks (MNNs)”, “Radial Basis Function Networks (RBFNs)”, “General Regression Neural Networks (GRNNs)” and “Recurrent Neural Networks (RNNs)”. These models are widely used in literature to forecast different attributes in the transport network. The five models used for comparative analysis were developed using the NeuralWorks Professional II software, a commercial package for neural network model development systems (NeuralWare, 2021). These models are well-established in the literature and the reader is referred to (Jacobs et al., 1991) for more details.

The parameters selected in this study and the learning rules and transfer function combinations that resulted in best performance are presented in Table 6.

3.2. Performance measures

Finding multiple suitable indicators for measuring prediction accuracy is essential, as no single indicator can measure the full precision/bias of predicted values. The following performance indicators were used to compare models (Vandeput, 2019). Therefore, for this analysis prediction accuracy is compared in three performance measures; MAPE, MAE and RMSE.

1. Mean Absolute Percentage Error (MAPE)

MAPE is a common key performance indicator used for prediction accuracy. MAPE divides each error by respective demand as in equation (7). In this measure, high errors during low-demand periods can significantly impact MAPE.

$$MAPE = \frac{1}{n} \sum \frac{|predicted\ value - demand|}{demand} \tag{7}$$

2. Mean Absolute Error (MAE) %

MAE is a good performance indicator also used to measure prediction accuracy. Key limitation is that it is not scaled to average demand. Therefore, MAE is divided by average passenger demand to obtain MAE % as in equation (8).

$$MAE\ \% = \frac{\sum |predicted\ value - demand|}{\sum demand} \tag{8}$$

3. Root Mean Squared Error (RMSE)

RMSE is a valuable key indicator calculated using Equation (9). RMSE is not scaled to the demand. Therefore, it is divided by average

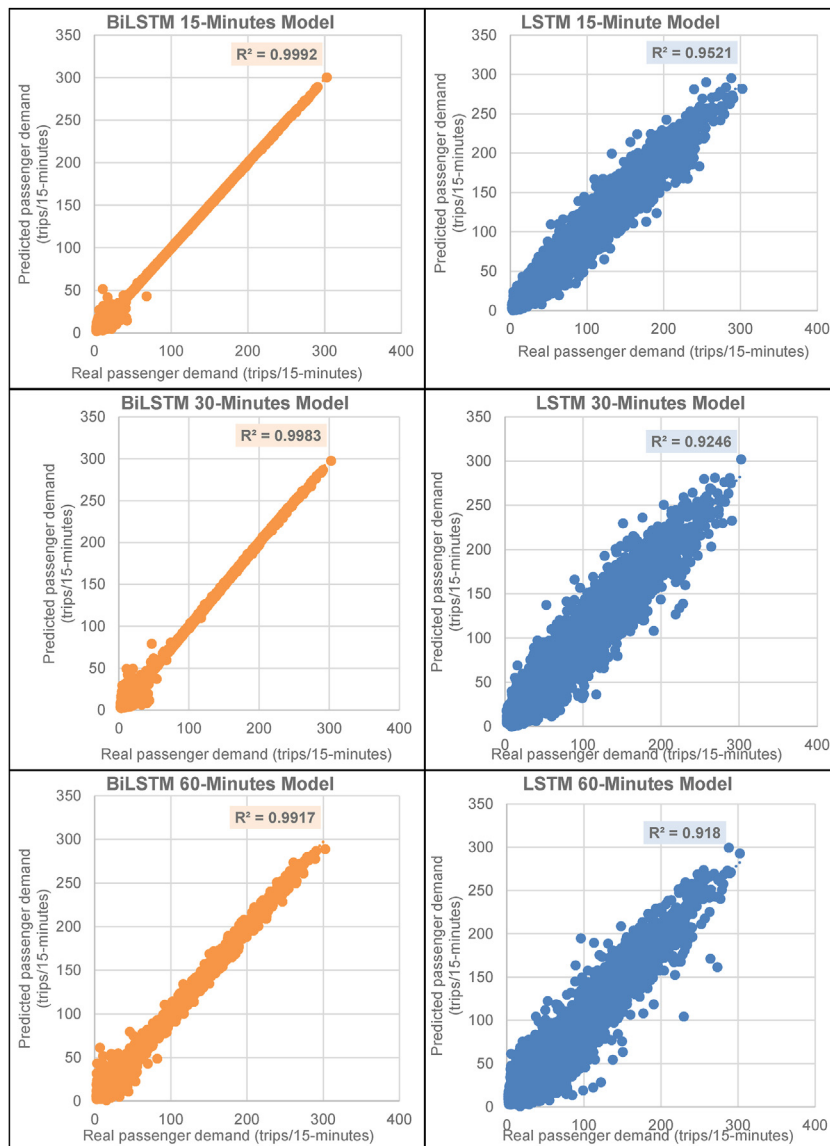


Fig. 6. Correlation between actual passenger demand to predicted passenger demand for BiLSTM and LSTM models.

demand to obtain RMSE %.

$$RMSE = \sqrt{\frac{1}{n} \sum (predicted\ value - demand)^2} \tag{9}$$

4. Short term prediction accuracies

This section summarises the results of passenger demand prediction for bus routes.

4.1. Model performance evaluation and validation

The results for demand prediction variations for bus routes are provided in Table 7. Here, green highlights represent the highest demand prediction accuracy, while yellow represents the second-best accuracy.

The BiLSTM models predicted bus demands with over 90% accuracy for 15-min and 30-min forecast horizons and approximately 80% accuracy for 1-h demand prediction. The results also show that the LSTM models, both uni-directional and bi-directional LSTM, have superior performance when compared to the other deep learning models such as RNN, GRNN, MNN, RBF, and DLBP models. For

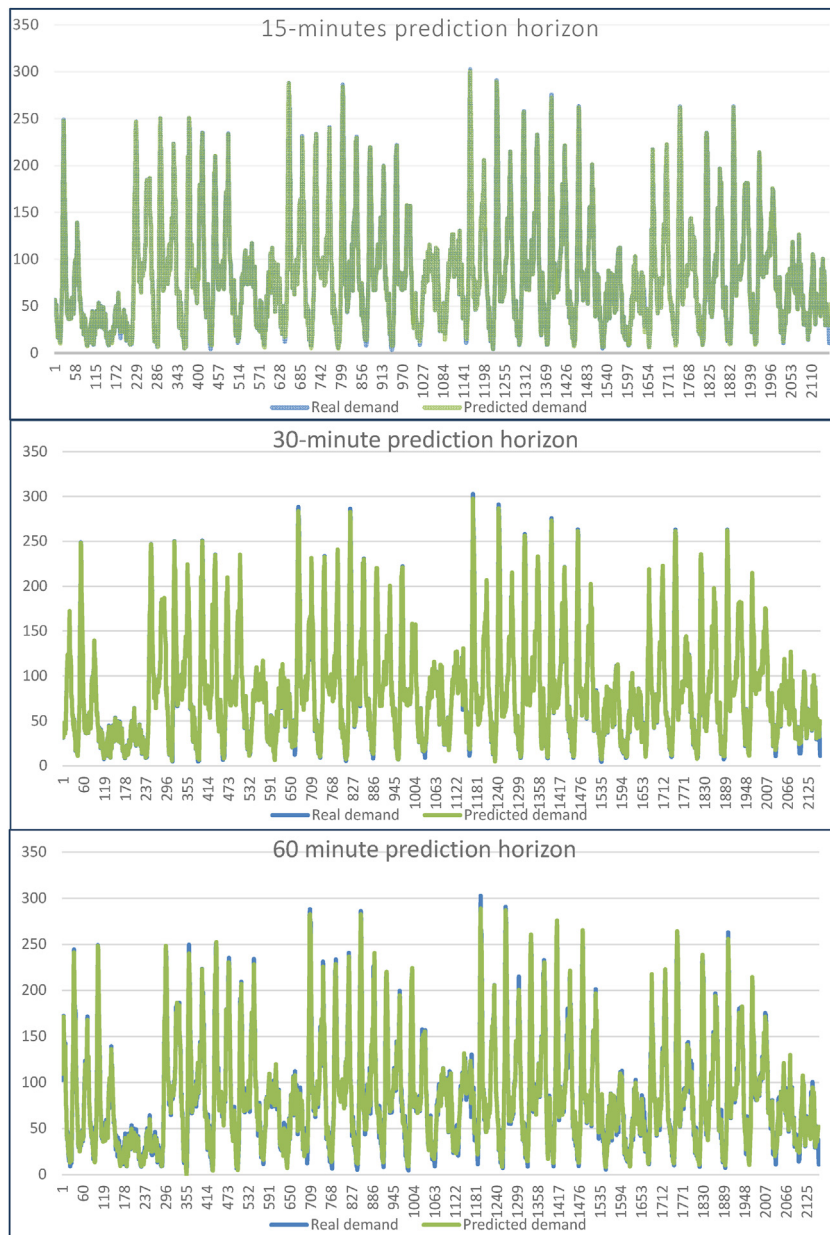


Fig. 7. Graphical representation of actual to predict passenger demand for BiLSTM models for (a) 15-min, (b) 30-min, and (c) 60-min time horizons.

bus route predictions, the results showed that BiLSTM provided a forecasting MAPE accuracy of around 97–99% for a 15-min prediction horizon, 94.77–98.40% for a 30-min prediction horizon, and 87.61–94.38% for hourly prediction. Compared to LSTM, the BiLSTM methodology shows the highest prediction accuracy, which does not deteriorate substantially with longer forecast horizons. The accuracy of the highest prediction horizon, 60-min, remained high at 87.61–94.38% demonstrating the capability to capture the complexity for the hourly prediction horizon.

4.2. Model fitting and prediction accuracy

Model fitting is measured with R-squared values. The BiLSTM model results showed a high correlation between actual and predicted demands as shown in Fig. 6. The LSTM model also showed over 0.9 R-squared values for 15, 30, and 60 min predictions. A comparison of the forecasting performance of the models is shown in Fig. 7 (a, b and c). These figures reflect the degree to which a model's prediction deviates from actual passenger demands on bus services. The hourly model has the lowest fitting degree, whereas the 15-min model shows the highest fitting degree.

Table 8
Comparisons of selected studies on bus passenger flow prediction.

Authors	Method	Contrast method	Method Style	Predict object	Data source	Data		Accuracy
						Temporal	Spatial	
Liyanae et al., 2021 (this study)	BiLSTM	DLBP, MNNS, RBFNs, GRNNs, RNNs	Single	Line	Smart-card	1-month	1,781 stops in 18- lines	MAPE 0.03–0.12 MAE 0.01–0.06 RMSE 0.03–0.08 Accuracy 92%–99%
Gong et al. (2014)	Kalman filter-based ARIMA	Direct-addition	Hybrid	Stop	APC and video	2-months	3-stops in 1-line	RE around 3%
Ma et al. (2014)	IMMPH with AR, SARIMA, ARIMA	ANNPH	Hybrid	Line	AFC	1-year	1 line	MAPE 5.82%
Zhai et al. (2020)	HTSDBNE	ARIMA (2,1,2) ELM, TS-ANN, SLMBP, SAE-DNN, MPDF	Hybrid	Stops	APC	6-months	2-stations in 2-lines	Error 8.787%
Xue et al. (2015)	IMM with ARMA, SARIMA, ARIMA	Real data	Hybrid	Line	AFC	4-months	1-line	MAPE 9.084%
Pekel and Soner Kara (2017)	POA-ANN IWD-ANN	GA-ANN	Hybrid	Line	AFC	NA	NA	MSE less 0.1
Zhang et al. (2017)	GM(1,1)	Real data	Single	Line	Manual survey	6-consecutive Mondays	1-line	RE less 10%
Bai et al. (2017)	MPDF	DBN, AE, AP-AE, FFNN, ARIMA	Hybrid	Line	Real data	5-months	1-line	MAPE 10.743%
Liu and Chen (2017)	SAE-DNN	Real data	Hybrid	Stops	AFC	4-months	4 stations in 7-lines	MAPE over 75%
Zhou et al. (2013)	Poisson model ARIMA	Real data	Hybrid	Stop	APTS	5-months	416 bus stops	Around 79%
Zhang et al. (2011)	Kalman filter	BP-ANN	Single	stop	AFC video	1-month	4-stations	Around 80%
Han et al. (2019)	LSTM Nadam-SGD	Naïve, ARIMA, SVR, 5 traditional LSTM	Hybrid	Stops	SCD	1-month	30-stations	MAPE 24.002%

5. Comparative evaluation with other studies in the literature

This section presents a comparative evaluation between the findings from this study and several similar studies reported in the literature. While previous sections of the paper compared different models using the same data, this section presents a holistic view of the effectiveness of our models compared to other literature. The comparisons in Table 8 are for three categories of models based on “traditional classical algorithms”, “regression models”, and “machine learning-based models” for single and hybrid models. The table presents the method used in each study, the comparisons conducted against other models, the architecture (single versus hybrid) used, the prediction object (route/line or bus stop), data source and dimensionality (temporal or spatial), and the accuracy of the model as reported in each paper.

The accuracy results clearly show that nonlinear methods performed better than linear methods. The hybrid models, which combine different time series methods to capture challenging data patterns and characteristics in their datasets, are highly complex and generally do not perform as well as single models due to the weakening of the universality of the hybrid models as a result of their more complex architectures. Among more recent literature, non-linear approaches based on machine learning seem to have been widely adopted for bus passenger demand prediction.

Establishing multiple suitable indicators for measuring prediction accuracy is essential, as no single indicator can measure the full precision and/or bias of predicted values. The performance indicators used to compare different models were based on existing literature (Vandeput, 2019) and included three key performance measures: MAPE, MAE and RMSE. As per Table 8, the BiLSTM model proposed in this study shows high prediction capability compared to the previous literature with an MAPE error of 0.03–0.12, MAE error in the range of 0.01–0.06 and RMSE with 0.03–0.08 range. With this, the prediction accuracy is over 92–99% for three different time horizons. This study therefore contributes to the body of knowledge in this field by using BiLSTM models with evidence of much higher accuracies exceeding 90%.

6. Conclusions and directions for future research

This paper presented robust deep learning models that were developed for short-term temporal predictions of passenger demand

using real-world data. The data was obtained from the MyKi smart-card fare payment system in Melbourne. Deep learning models were constructed representing 15-min, 30-min, and 60-min. These models were developed using BiLSTM and LSTM deep learning methodologies based on one month of data comprising 27,823 data points for the 15-min model, 27,360 data points for the 30-min model, and 26,561 data points for the 60-min model. The findings of this study showed that both the BiLSTM and LSTM architectures provided the highest predictive intelligence accuracy of over 90% for short-term predictions of passenger demands for 15-min, 30-min and 60-min time horizons.

The main limitation of this study was limited access to smart-card data which in our case was constrained to one month of data between 1–27 May 2018. This was the only data made available to the researchers. Access to larger data sets covering more months and years as well as more routes can help improve the accuracy and reliability of the models even further. Furthermore, for this analysis, aggregated passenger demand for inbound and outbound direction for each route was necessary because the data did not have sufficient observations to enable separate analyses for inbound and outbound directions. It is noted here, however, that the direction of the service will not have an impact on model selection and performance. The key factor influencing performance, in either inbound or outbound directions, is the availability of quality representative data that can be used for model training and testing. We aim to address this limitation in future studies through analysis of large data sets that can provide accurate and sufficient observations in both directions of travel.

The AI-inspired deep learning passenger demand forecasting models can be applied in a wide variety of public transport operational contexts. In this work, we estimated future passenger demands based on routes. The same methodology can be applied to estimate passenger demands at bus and tram stops and also train stations where this information can be used to predict the number of buses, trams or even number of compartments on a train that are needed to meet passenger demands for different routes at different times of the day. Knowing this information in advance can save operators money through reducing costs of operations. It can also improve customer satisfaction and enhance the sustainability outcomes for urban areas through promoting reliable public transport modes of urban mobility and shifting drivers away from energy-intensive private vehicles to more energy-efficient modes of public transport.

Researchers interested in extending this work should also consider improving the performance and practicality of Bi-LSTM models for more data structures, such as daily, weekly, and monthly aggregations based on different aspects including route, service direction, and at points of boarding and disembarking such as bus and tram stops and at train stations. Future studies should also consider more patterns representing the impacts of particular weather conditions and traffic incidents. Other aspects of extending this research would be considering advanced deep learning architectures and the influence of network parameters on model performance.

Ethics clearance

This work was carried out in accordance with Swinburne University's Human Research Ethics Guidelines. Project ref: 20214391–6022.

Acknowledgements of Research Funding

Sohani Liyanage and Rusul Abduljabbar acknowledge the Swinburne University of Technology for their PhD scholarships. Rusul Abduljabbar also acknowledges her scholarship from the Iraqi Government. The authors acknowledge the Department of Transport, Victoria, for providing one month of MyKi data for use in this research.

References

- Abduljabbar, R. L., Dia, H., & Tsai, P.-W. (2021). Unidirectional and bidirectional LSTM models for short-term traffic prediction. *Journal of Advanced Transportation*, 2021.
- Abduljabbar, R. L., Dia, H., Tsai, P.-W., & Liyanage, S. (2021). Short-term traffic forecasting: An LSTM network for spatial-temporal speed prediction. *FutureTransportation*, 1, 21–37.
- Adams, D., Novak, A., Klietnik, T., & Potcovaru, A.-M. (2021). Sensor-based big data applications and environmentally sustainable urban development in internet of things-enabled smart cities. *Geopolitics, History, and International Relations*, 13, 108–118.
- Aldridge, S., & Stehel, V. (2021). Intelligent vehicular networks, deep learning-based sensing technologies, and big data-driven algorithmic decision-making in smart transportation systems. *Contemporary Readings in Law and Social Justice*, 13, 107–120.
- Anvari, S., Tuna, S., Canci, M., & Turkay, M. (2016). Automated Box–Jenkins forecasting tool with an application for passenger demand in urban rail systems. *Journal of Advanced Transportation*, 50, 25–49.
- Bai, Y., Sun, Z., Zeng, B., Deng, J., & Li, C. (2017). A multi-pattern deep fusion model for short-term bus passenger flow forecasting. *Applied Soft Computing*, 58, 669–680.
- Balcombe, R. (2004). *The demand for public transport: A practical guide, TRL report, TRL593*. TRL Limited.
- Banerjee, N., Morton, A., & Akartunali, K. (2020). Passenger demand forecasting in scheduled transportation. *European Journal of Operational Research*, 286, 797–810. <https://doi.org/10.1016/j.ejor.2019.10.032>
- Bengio, Y. (2009). *Learning deep architectures for AI*. Now Publishers Inc.
- Bengio, Y., Lamblin, P., Popovici, D., & Larochelle, H. (2007). Greedy layer-wise training of deep networks. In *Advances in neural information processing systems* (pp. 153–160).
- Blake, R., & Michalikova, K. F. (2021). Deep learning-based sensing technologies, artificial intelligence-based decision-making algorithms, and big geospatial data analytics in cognitive internet of things. *Analysis and Metaphysics*, 20, 159–173.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. John Wiley & Sons.
- Buckley, J. J., & Hayashi, Y. (1994). Fuzzy neural networks: A survey. *Fuzzy sets Syst*, 66, 1–13.
- Ceder, A., Hassold, S., Dunlop, C., & Chen, I. (2013). Improving urban public transport service using new timetabling strategies with different vehicle sizes. *International Journal of Urban Science*, 17, 239–258.
- Chen, R., Liang, C.-Y., Hong, W.-C., & Gu, D.-X. (2015). Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm. *Applied Soft Computing*, 26, 435–443.

- Chen, Y., Yang, B., & Meng, Q. (2012). Small-time scale network traffic prediction based on flexible neural tree. *Applied Soft Computing*, 12, 274–279.
- Chen, Z., Zhao, B., Wang, Y., Duan, Z., & Zhao, X. (2020). Multitask learning and GCN-based taxi demand prediction for a traffic road network. *Sensors*, 20, 3776.
- Cyril, A., Mulangi, R. H., & George, V. (2018). Modelling and forecasting bus passenger demand using time series method. In *2018 7th international conference on reliability, infocom technologies and optimization (trends and future directions)(ICRITO)* (pp. 460–466). IEEE.
- Dia, H. (2001). An object-oriented neural network approach to short-term traffic forecasting. *European Journal of Operational Research*, 131, 253–261.
- Dia, H., Harney, D., & Boyle, A. (2001). Dynamics of drivers' route choice decisions under advanced traveller information systems. *Road and Transport Research*, 10, 3.
- Dia, H., & Panwai, S. (2011). Neural agent (Neugent) models of driver behavior for supporting ITS simulations. *International Journal of Intelligence Transport System Research*, 9, 23–36.
- Dia, H., & Rose, G. (1995). Development of artificial neural network models for automated detection of freeway incidents. In *7th world conference on transport research*. Citeseer.
- Dia, H., Rose, G., & Snell, A. (1996). *Comparative performance of freeway automated incident detection algorithms*.
- Dou, F., Xu, J., Wang, L., & Jia, L. (2013). A train dispatching model based on fuzzy passenger demand forecasting during holidays. *Journal of Industrial Engineering and Management*, 6, 320–335.
- Faridai, S., Juraveva, R. S., Darovskikh, S. N., & Sh, Q. S. (2021). Neural network model for predicting passenger congestion to optimize traffic management for urban public transport. *Computational Technology Management Electron*, 21.
- Gan, M., Cheng, Y., Liu, K., & Zhang, G. (2014). Seasonal and trend time series forecasting based on a quasi-linear autoregressive model. *Applied Soft Computing*, 24, 13–18.
- Gong, M., Fei, X., Wang, Z. H., & Qiu, Y. J. (2014). Sequential framework for short-term passenger flow prediction at bus stop. *Transportation Research Record*, 2417, 58–66.
- Guo, S., Li, W., Bai, W., & Zhang, D. (2013). Prediction of short-term passenger flow on a bus stop based on LSSVM. *Journal of Wuhan University Technology (Transportation Science Engineering)*, 37, 603–607.
- Hai-lan, L., & En-chong, X. I. (2012). Prediction model of public transport passenger volume based on ran-dom fluctuations. *Journal Chang University (Natural Science Ed.)*, 1.
- Han, Y., Wang, C., Ren, Y., Wang, S., Zheng, H., & Chen, G. (2019). Short-term prediction of bus passenger flow based on a hybrid optimized LSTM network. *ISPRS International Journal of Geo-Information*, 8, 366.
- He, Y., Zhao, Y., & Tsui, K.-L. (2022). Short-term forecasting of origin-destination matrix in transit system via a deep learning approach. *Transport A Transport Science*, 1–28. <https://doi.org/10.1080/23249935.2022.2033348>
- Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18, 1527–1554.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9, 1735–1780.
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural Computation*, 3, 79–87.
- Jiang, X., Zhang, L., & Chen, X. M. (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: Emerging Technologies*, 44, 110–127.
- Ke, J., Qin, X., Yang, H., Zheng, Z., Zhu, Z., & Ye, J. (2021). Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network. *Transportation Research Part C: Emerging Technologies*, 122, Article 102858.
- Ke, J., Zheng, H., Yang, H., & Chen, X. M. (2017). Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. *Transportation Research Part C: Emerging Technologies*, 85, 591–608.
- Lăzăroi, G., & Harrison, A. (2021). Internet of things sensing infrastructures and data-driven planning technologies in smart sustainable city governance and management. *Geopolitics, History, and International Relations*, 13.
- Lee, S., Lee, Y., & Cho, B. (2006). Short-term travel speed prediction models in car navigation systems. *Journal of Advanced Transportation*, 40, 122–139.
- Liang, X., Wang, G., Min, M. R., Qi, Y., & Han, Z. (2019). A deep spatio-temporal fuzzy neural network for passenger demand prediction. In *Proceedings of the 2019 SIAM international conference on data mining* (pp. 100–108). SIAM.
- Li, C., Bai, L., Liu, W., Yao, L., & Waller, S. T. (2020). Knowledge adaption for demand prediction based on multi-task memory neural network. In *Proceedings of the 29th ACM international conference on information & knowledge management* (pp. 715–724).
- Lí, C., Bai, Y., & Zeng, B. (2016). Deep feature learning architectures for daily reservoir inflow forecasting. *Water Resour. OR Manager*, 30, 5145–5161.
- Li, C., Cabrera, D., De Oliveira, J. V., Sanchez, R.-V., Cerrada, M., & Zurita, G. (2016). Extracting repetitive transients for rotating machinery diagnosis using multiscale clustered grey infogram. *Mechanical Systems and Signal Processing*, 76, 157–173.
- Lí, D., Cao, J., Lí, R., & Wu, L. (2020). A spatio-temporal structured lstm model for short-term prediction of origin-destination matrix in rail transit with multisource data. *IEEE Access*, 8, 84000–84019.
- Liu, L., & Chen, R.-C. (2017). A novel passenger flow prediction model using deep learning methods. *Transportation Research Part C: Emerging Technologies*, 84, 74–91.
- Liu, Y., Liu, Z., & Jia, R. (2019). DeepPF: A deep learning based architecture for metro passenger flow prediction. *Transportation Research Part C: Emerging Technologies*, 101, 18–34.
- Liu, W., Tan, Q., & Wu, W. (2020). Forecast and early warning of regional bus passenger flow based on machine learning. *Mathematical Problems in Engineering*, Article 6625435. <https://doi.org/10.1155/2020/6625435>, 2020.
- Liu, G., Yin, Z., Jia, Y., & Xie, Y. (2017). Passenger flow estimation based on convolutional neural network in public transportation system. *Knowledge-Based Systems*, 123, 102–115.
- Liyanage, S., & Dia, H. (2020). An agent-based simulation approach for evaluating the performance of on-demand bus services. *Sustainability*, 12, 4117.
- Liyanage, S., Dia, H., Abduljabbar, R., & Bagloe, S. A. (2019). Flexible mobility on-demand: An environmental scan. *Sustainable Times*, 11. <https://doi.org/10.3390/su11051262>
- Lu, X., Ma, C., & Qiao, Y. (2021). Short-term demand forecasting for online car-hailing using ConvLSTM networks. *Physics A Statistics Mechanical its Application*, 570, Article 125838. <https://doi.org/10.1016/j.physa.2021.125838>
- Matlab. (2021). *Long short-term memory networks [WWW document]*. URL: Matlab, 2.9.21 <https://au.mathworks.com/help/deeplearning/ug/long-shortterm-0Amemory-networks.html>.
- Ma, Z., Xing, J., Mesbah, M., & Ferreira, L. (2014). Predicting short-term bus passenger demand using a pattern hybrid approach. *Transportation Research Part C: Emerging Technologies*, 39, 148–163.
- Milenković, M., Švadlenka, L., Melichar, V., Bojović, N., & Avramović, Z. (2018). SARIMA modelling approach for railway passenger flow forecasting. *Transport*, 33, 1113–1120.
- NeuralWare. (2021). *NeuralWare neural computing, using NeuralWorks, and reference GuideNeuralWare [WWW document]*. URL: Pittsburgh, PA www.neuralware.com, 2.9.21.
- Nica, E. (2021). Urban Big Data analytics and sustainable governance networks in integrated smart city planning and management. *Geopolitics, History, and International Relations*, 13, 93–106.
- Noekel, K. (2016). *Modelling public transport passenger flows in the era of intelligent transport systems: COST action TU1004 (TransITS)*. Springer.
- Panwai, S., & Dia, H. (2005). Development and evaluation of a reactive agent-based car following model. In *Proceedings of the intelligent vehicles and road infrastructure conference*. Citeseer.
- Panwai, S., & Dia, H. F. (2006). *A fuzzy neural approach to modelling behavioural rules in agent-based route choice models*.
- Pekel, E., & Soner Kara, S. (2017). Passenger flow prediction based on newly adopted algorithms. *Applied Artificial Intelligence*, 31, 64–79.
- Pelau, C., Dabija, D.-C., & Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Computers in Human Behavior*, 122, Article 106855.
- PTV. (2019). *Timetables - public transport Victoria [WWW document]*. PTV. URL <https://www.ptv.vic.gov.au/timetables>.

- Roos, J., Gavin, G., & Bonnevey, S. (2017). A dynamic Bayesian network approach to forecast short-term urban rail passenger flows with incomplete data. *Transport Respository procedia*, 26, 53–61.
- Rose, G., & Dia, H. (1995). Freeway automatic incident detection using artificial neural networks. In *Proceedings of the international conference on application of new Technology to transport systems* (pp. 123–140). Citeseer.
- Rowland, Z., & Porter, K. (2021). Autonomous vehicle driving algorithms, deep learning-based sensing technologies, and big geospatial data analytics in smart sustainable intelligent transportation systems. *Contemporary Readings in Law and Social Justice*, 13, 23–36.
- Smith, B. L., Williams, B. M., & Oswald, R. K. (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 10, 303–321.
- 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, S., Zhang, C., & Yu, G. (2006). A Bayesian network approach to traffic flow forecasting. *IEEE Transactions on Intelligent Transportation Systems*, 7, 124–132.
- Thomas, K., & Dia, H. (2006). Comparative evaluation of freeway incident detection models using field data. In *IEE proceedings-intelligent transport systems* (pp. 230–241). IET.
- Thomas, K., Dia, H., & Cottman, N. (2001). Simulation of arterial incident detection using neural networks. In *Proceedings of the 8th world congress on intelligent transport systems*.
- Tirachini, A., Hensher, D. A., & Rose, J. M. (2013). Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. *Transport Respository Part A Policy Pract.*, 53, 36–52.
- Toqué, F., Côme, E., El Mahrsi, M. K., & Oukhellou, L. (2016). Forecasting dynamic public transport origin-destination matrices with long-short term memory recurrent neural networks. In *2016 IEEE 19th international conference on intelligent transportation systems (ITSC)* (pp. 1071–1076). IEEE.
- 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, 3728–3736.
- Vandeput, N. (2019). *Forecast KPIs: RMSE, MAE, MAPE & bias*. Data Science for Supply Chain Forecasting.
- Vlahogianni, E., & Karlaftis, M. (2011). Temporal aggregation in traffic data: Implications for statistical characteristics and model choice. *Transport Letter*, 3, 37–49.
- Wallace, S., & Lăzăroi, G. (2021). Predictive control algorithms, real-world connected vehicle data, and smart mobility technologies in intelligent transportation planning and engineering. *Contemporary Readings in Law and Social Justice*, 13, 79–92.
- Wang, Y., Yin, H., Chen, H., Wo, T., Xu, J., & Zheng, K. (2019). Origin-destination matrix prediction via graph convolution: A new perspective of passenger demand modeling. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 1227–1235).
- Wang, Q., & Zhang, Q. (2012). Forecasting of short-term urban public transit volume based on random gray ant colony neural network. *Application Research of Computers*, 29, 2078–2080.
- Wei, Y., & Chen, M.-C. (2012). Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. *Transportation Research Part C: Emerging Technologies*, 21, 148–162.
- World population review. (2021). World population review [WWW document], 11.7.21, URL <https://worldpopulationreview.com/world-cities/melbourne-population>.**
- Wu, W., Jiang, S., Liu, R., Jin, W., & Ma, C. (2020). Economic development, demographic characteristics, road network and traffic accidents in zhongshan, China: Gradient boosting decision tree model. *Transport A Transport Science*, 16, 359–387.
- Wu, W., Li, P., Liu, R., Jin, W., Yao, B., Xie, Y., & Ma, C. (2020). Predicting peak load of bus routes with supply optimization and scaled shepard interpolation: A news vendor model. *Transport Respository Part E Logist. Transp. Rev.*, 142, Article 102041.
- Wu, W., Xia, Y., & Jin, W. (2020). Predicting bus passenger flow and prioritizing influential factors using multi-source data: Scaled stacking gradient boosting decision trees. *IEEE Transactions on Intelligent Transportation Systems*, 22, 2510–2523.
- Xue, R., Sun, D. J., & Chen, S. (2015). Short-term bus passenger demand prediction based on time series model and interactive multiple model approach. *Discrete Dynamical Natural Society* 2015.
- Yang, X., & Liu, L. (2016). Short-term passenger flow forecasting on bus station based on affinity propagation and support vector machine. *Journal of Wuhan University Technology (Transportation Science Engineering)*, 40, 36–40.
- Yang, Q., Yang, Y. F., Feng, Z. X., & Zhao, X. W. (2013). Prediction method for passenger volume of city public transit based on grey theory and Markov model. *China Journal of Highway and Transport*, 26, 169–175.
- Yeon, K., Min, K., Shin, J., Sunwoo, M., & Han, M. (2019). Ego-vehicle speed prediction using a long short-term memory based recurrent neural network. *International Journal of Automotive Technology*, 20, 713–722.
- Zhai, H., Cui, L., Nie, Y., Xu, X., & Zhang, W. (2018). A comprehensive comparative analysis of the basic theory of the short term bus passenger flow prediction. *Symmetry*, 10, 369.
- Zhai, H., Tian, R., Cui, L., Xu, X., & Zhang, W. (2020). A novel hierarchical hybrid model for short-term bus passenger flow forecasting. *Journal of Advanced Transport* 2020.
- Zhang, J., Che, H., Chen, F., Ma, W., & He, Z. (2021). Short-term origin-destination demand prediction in urban rail transit systems: A channel-wise attentive split-convolutional neural network method. *Transportation Research Part C: Emerging Technologies*, 124, Article 102928.
- Zhang, K., Liu, Z., & Zheng, L. (2019). Short-term prediction of passenger demand in multi-zone level: Temporal convolutional neural network with multi-task learning. *IEEE Transactions on Intelligent Transportation Systems*, 21, 1480–1490.
- Zhang, X., Mao, B., Wang, Y., Feng, J., & Li, M. (2013). Wavelet neural network-based short-term passenger flow forecasting on urban rail transit. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 11, 7379–7385.
- Zhang, C.-H., Song, R., & Sun, Y. (2011). Kalman filter-based short-term passenger flow forecasting on bus stop. *Journal of Transport System Engineering Information Technology*, 11, 154–159.
- Zhang, D., Xiao, F., Shen, M., & Zhong, S. (2021). Dneat: A novel dynamic node-edge attention network for origin-destination demand prediction. *Transportation Research Part C: Emerging Technologies*, 122, Article 102851.
- Zhang, Z., Xu, X., & Wang, Z. (2017). Application of grey prediction model to short-time passenger flow forecast. In *AIP conference proceedings* (p. 20136). AIP Publishing LLC.
- Zhou, C., Dai, P., & Li, R. (2013). The passenger demand prediction model on bus networks. In *2013 IEEE 13th international conference on data mining workshops* (pp. 1069–1076). IEEE.