Integrating Bluetooth and smart card data for better estimation and prediction of bus speed on arterial corridors with low frequency buses

Faria Shanjana Imam[1], Ashish Bhaskar[1], Edward Chung[1]

[1]Smart Transport Research Centre, School of Civil Engineering and Built Environment, Queensland University of Technology, Brisbane, Australia

Email for correspondence: fariashanjana.imam@hdr.qut.edu.au

Abstract

Data driven based travel speed (or travel time) short term prediction models require accurate estimation of the historical time series with equally spaced data points. The availability of the bus speed time series data points depends on the bus frequency and other operational factors such as on-time performance. Low frequency bus routes, coupled with bad on-time performance can result in time series with number of missing values (or irregular interval of data points).

Addressing the above need, this paper explores the relationship between bus and car speed and utilises the understanding to better estimate bus travel speed time series and its application for short term prediction of bus speed. With a case study on a Brisbane corridor, the car speed is estimated using Bluetooth MAC Scanner (BMS) and bus speed is estimated using Automatic Fare Collection data (Go card). The findings are encouraging and the results of the integration of the two data sources indicate around 3% improvement in the bus speed estimation compared to the case where the time series gaps are filled with linear interpolation. Furthermore, the prediction results are also improved for different prediction horizons.

This paper will assist transit operators to exploit Bluetooth data to augment the performance of low frequency buses by estimating and predicting more accurate bus speed (travel time).

Keywords: Automatic Fare Collection, Bluetooth, Bus speed, Car speed, Travel time, urban corridor, estimation, prediction

1 Introduction

Travel time information is an important variable for both travellers' decisions related to mode choice and route choice and operator's management and control of the transport network. The desired advanced traveller information system should be timely, accurate and reliable. Research on travel time or travel speed has been done for long. Though most of this research is limited for freeways(Bhaskar et al., 2014a) research on arterial networks (Bhaskar et al., 2014b, Khoei et al., 2013) has also gained interest since last decade. There are differences between freeway and arterial which makes the travel time estimation and prediction more complicated for arterial road. This includes traffic signal, bus stops, and pedestrian crossings on the arterial network.

In literature, different models for bus (Chien and Kuchipudi 2003; Mazloumi et al. 2011) and car (Fei, Lu and Liu 2011; Hinsbergen, Lint and Zuylen 2009)(Fadaei et al., 2017) average travel time predictions are proposed. Most of the data driven based models requires estimation of the time series of travel time. Time series is a series of data points which are equally spaced in time. For accurate prediction of travel time, one need to first accurately estimate the database on time series of travel time. The estimation of the time series is challenging for buses, due to lack of the availability of data points at equally spaced time intervals. For instance, if one is interested to define bus travel speed time series at interval of 5 minutes, then scheduled headway between buses should be at least 5 minutes. However, due to disturbances in bus network, there is irregularity in the bus scheduled on-time performance which leads to intervals in time series where there is no data available. The problem is more evident when the headway between buses is higher. Generally, the gaps in the time series are filled by linear interpolation which might not be accurate during congested conditions.

With advancement in technology, rich traffic data from arterial networks are available. This includes, Bluetooth MAC Scanners (BMS) (Bhaskar and Chung, 2013), Smart Card (automatic fare collection for public transit), and Automatic Vehicle Location (AVL).The availability of such data provides avenues for estimation and prediction of travel time (travel speed) for both car and buses on the arterial networks.

This research focuses on data driven based bus travel speed (travel time) estimation and short term prediction for arterial corridors where bus and car share the road. We use the term travel speed and travel time interchangeable, because speed over a corridor is the ratio of the length of the corridor and the travel time over the corridor. Here, estimation is the development of the historical database for time series of the travel speed and short term prediction is forecasting the bus speed up to 30 minutes in future. With a case study on the real data from Brisbane, the objective of this paper is to 1) Model time series of bus speed by integrating travel speed from bus and car and 2) Evaluate the accuracy of the estimated bus time series for short term travel speed prediction.

Here, bus speed is obtained from the smart card data. BMS scans the MAC ID of the electronic devices that can be from any mode of travel (such as car and bus). Bhaskar et al., (2015) has empirically analysed Brisbane BMS and bus Vehicle Identification and Detection data (VID) and has observed that bus is not overrepresented in the BMS data. The research focuses on the corridors with low frequency buses, which further reduces the chances of bus representation in the BMS data. Therefore, the filtered BMS data can be assumed to be a good representative of average car travel speed.

The remaining of the paper is structured as follows. Section 2 provides insight on the literature related to the topic. Section **Error! Reference source not found.** introduces the study site and the data used for this research. Thereafter, section 4 describes research methodology. Finally, the model is cross-validated and the paper is concluded after presenting the results of analysis in section 5.

2 Literature review

Many different models have been developed for travel time prediction or estimation including Time series methods, Kalman Filtering, Nonparametric Regression Models and Neural Networks for both bus and car for freeway and arterial. Though most of the literatures are found on car travel time estimation and prediction (Bhaskar, Chung and Dumont 2011, 2010; Fei, Lu and Liu 2011), a good number of studies has also been found on bus travel time prediction. In most of the cases bus and car travel time are estimated or predicted separately thus creating a gap in the literature for a transferable multimodal travel time relationship for large scale application. Few studies used buses as traffic probes in urban areas (Bertini and Tantiyanugulchai 2004; Cathey and Dailey 2001) to estimate the car travel time or speed. The TriMet study (Bertini and Tantiyanugulchai 2004) applied a reverse linear regression technique and found the following relationship existed:

$$CTS = 0.72 * MIBS + \varepsilon \qquad (1)$$

Where, MIBS= maximum instantaneous bus speed between two adjacent stops
ε = random error

The following relationship between the bus travel time and car travel time in Blacksburg was proposed by Bae (1995).

$$y = 0.429459842 + 0.665057187*x \qquad (2)$$

where, y = car travel time
x = bus travel time
Although these empirical equations are site-specific it does provide the good and concise indication of travel time correlations between two different major modes of transportation.

Levinson (1983) analysed the traffic data from few U.S cities and revealed that average car speeds are 1.4 to 1.6 times faster than average bus speed. McKnight et al. (2004) proposed an equation between bus travel time per mile and car travel time per mile. Kieu et al., (2012) developed a relationship between bus and car travel time on urban networks by utilising the empirical Bluetooth and Bus Vehicle Identification data from Brisbane. The results revealed that the not-in-service bus travel time are similar to the car travel time and the in-service bus travel time could be used to estimate car travel time during off-peak hours. Kieu et al., (2015) enumerated the relationship between bus speed and average car speed by integrating Bluetooth data and Transit Signal Priority data from the urban network in Brisbane, Australia. Nevertheless, this study did not aim to estimate real-time bus speed using car speed and vice versa.

Most of these studies concluded that buses could be used as traffic probes effectively in urban arterials. To the author's knowledge no study aimed to get the benefit of cars on a shared road for the better estimation and prediction of bus speed. As discussed in the introduction section, the knowledge of the car speed should be useful to enhance the estimation of the time-series of bus speed for corridors with low frequency of buses. This paper fills this gap.

3 Research hypothesis

In this research, we make a hypothesis that the availability of the car speed should contribute to the improvement of bus speed estimation for corridors with low frequency of buses. To test the hypothesis we integrate AFC data with BMS data and develop a relationship between the bus and car speed. Thereafter, the missing values in the time series of the bus speed is estimated using the relationship.
The benefits of the accurate estimation of the bus time series is evaluated through its application for short term prediction of the bus travel speed. For the short term speed prediction this study uses Artificial Neural Network (ANN) modelling. In literature, ANN is acknowledged for its capability to solve complex non-linear relationships and is widely used in travel time prediction studies (Jeong and Rilett 2004; Chien, Ding and Wei 2002; Mazloumi et al. 2011; Lin et al. 2013).

4 Study site and data

The research uses real data from Brisbane. It includes BMS data from Brisbane City Council and AFC data (smart card named Go Card) from Translink (a division of Queensland Department of Transport and Main Roads). Modelling speed is equivalent to modelling travel time normalised with distance. As data from different sources are used, speed is chosen as the direct parameter for modelling in this paper.

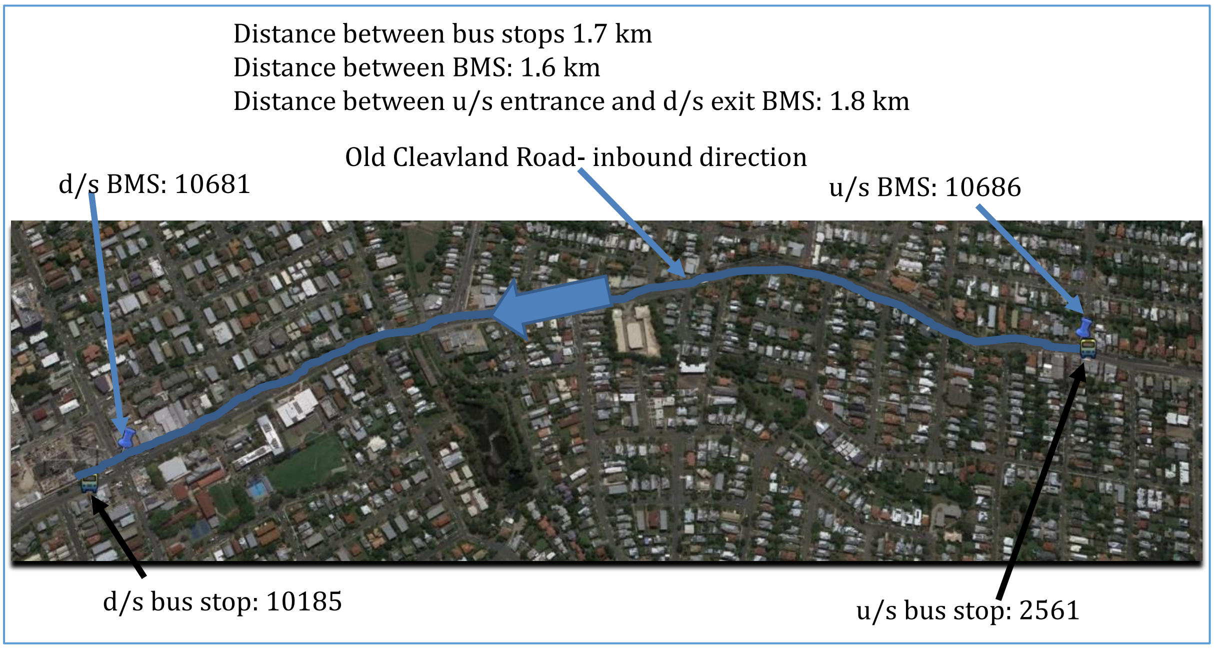
We are interested in the study site with:

a) high frequency of buses so that the ground truth time series of bus speed can be estimated with confidence;
b) the availability of car speed through BMS scanners;
c) Bus and car share the road, i.e. no exclusive bus lane; and
d) Traffic is sufficiently high and the corridor is congested during peak period

The above criteria are satisfied for around 2 km of the inbound traffic (towards City Centre) on Old Cleveland Road, Brisbane. **Figure *1*** illustrates the study section, where BMS is presented as blue markers and bus stops are presented with a 'bus' marker. We define our study corridor from upstream bus stop 2561 to downstream bus stop number 10185. These stops are served by seven bus routes with over 42 buses during the peak period. The distance between these bus stops is 1.7 km. Between the above upstream and downstream bus stops there are additional five bus stops and four signalised intersection. The speed limit of the section is 60 km/hr.

The nearby BMS for upstream and downstream bus stop is BMS 10686 and 10681, respectively. The distance between BMS scanners is 1.6 km. As can be seen from **Figure *1***, the stops are closer to the BMS. Brisbane City Council (BCC) BMS has a scanning zone of around 100 m. Refer to Bhaskar and Chung (2013) for the detailed understanding of the BMS. The stop at u/s is slightly upstream of the BMS and the stop at d/s is around 50 m downstream of the d/s BMS. Therefore, for this study the car speed is estimated considering the time when the Bluetooth MAC Id is observed at the entrance of the upstream BMS and exit of the downstream BMS. The distance between the u/s entrance and d/s exit of the BMSs is 1.8 km.

Figure 1: Study site, Old Cleveland Road, Brisbane



Data from working days (excluding school holidays) from May 2015 to March 2016 (10 months) is used. 70% of the data is used for model development and calibration. Remaining 30 % is used for model validation. For the current analysis only the morning peak period (7:00 am to 8:30 am) is considered.

5 Methodology

The individual car and bus travel times are estimated from Bluetooth and Go card data, respectively. Individual vehicle travel times are cleansed by applying filters and travel time are transformed into individual vehicle space mean speed by considering the respective corridor length. For Bluetooth data, we have applied the filter of Median Absolute Deviation (MAD). Refer to Bhaskar et al., (2014b) for the details of the filter.

The travel time obtained from both BMS and Go card data is transformed into speed using the respective distance between u/s and d/s scanners and stops. This provides a time series of individual vehicle car speeds ($v_{car,i}$) and bus speeds ($v_{bus,j}$) corresponding to the time when the vehicle was observed at upstream. Here, *i* and *j* are number of cars and buses observed at upstream at time *t*, respectively.

For each time series, we define a time window of 5 minutes [t-2.5 minutes, t+2.5 minutes] and estimate the time series for the (5 minutes) average bus speed (v_{bus}) and average car speed (v_{car}).

We develop the following bus time series:

a) Ground truth **(TS_G)**: Here all the bus routes are considered. This series is used to validate the estimation and prediction models.
b) Time series from low frequency buses: Here we consider a single route only. This is to replicate the scenario when there is low frequency bus route. This series has number of missing values. For testing the hypothesis, we consider filling the missing values with
    a. Linear interpolation (base case): This provides a time series where the missing values are obtained through linear interpolation **(TS_L)**
    b. An estimate from car speed (proposed case): This provides a time series where the missing values are obtained through knowledge from car speed **(TS_C).**

5.1 BMS based time series

BMS data includes

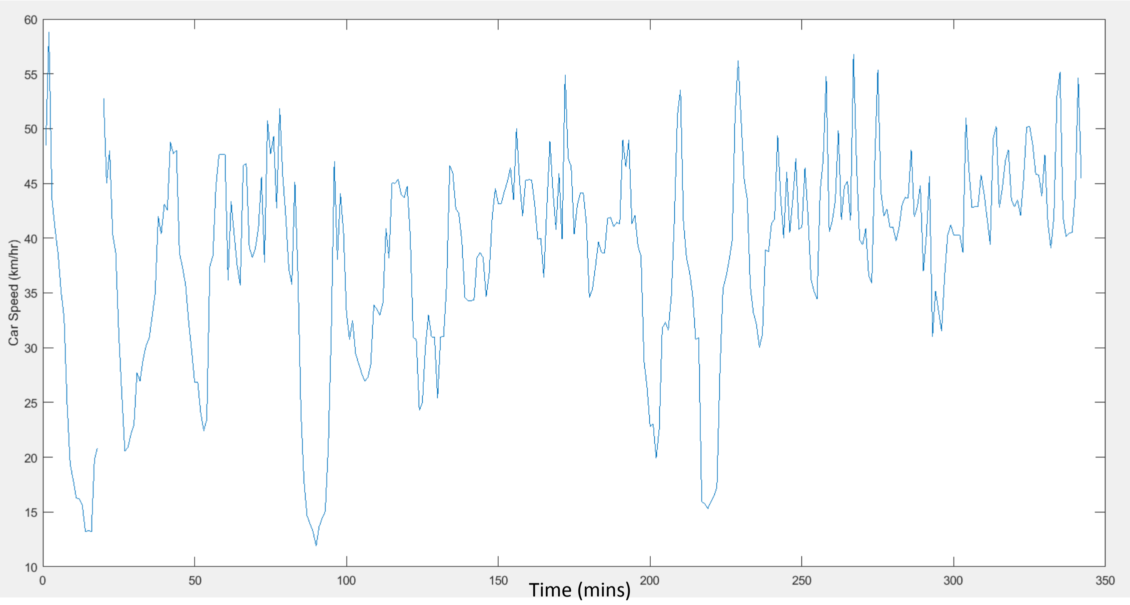
a) *Location ID:* ID where the scanner is located
b) *MACID:* The observed MAC ID
c) *Timestamp:* Time when the MAC ID is first observed within the scanning location
d) *Duration:* Time gap between the first and the last observation of the MAC ID at the location

Here, travel time, *TT(m,u,d)*, of a Bluetooth MAC-ID , *m*, observation at upstream, *u*, and downstream, *d*, is defining as the difference of the time when the device is observed at downstream, *T(m,d)*, and upstream, *T(m,u)*, scanners. Refer to Tsubota et al., (Tsubota et al., 2011) for the Bluetooth travel time estimation methodology.

$$TT(m,u,d)=TT(m,d)-TT(m,u) \quad (3)$$

An example of a time series for car speed from BMS is presented in **Figure 2**.

Figure 2: Sample car speed from BCC



5.2 AFC based time series

The bus travel time is obtained from AFC data. In Brisbane, the passengers need to *touch on* and *touch off* the smart card each time they board and alight the bus, respectively. The data includes the following key fields used in this research:

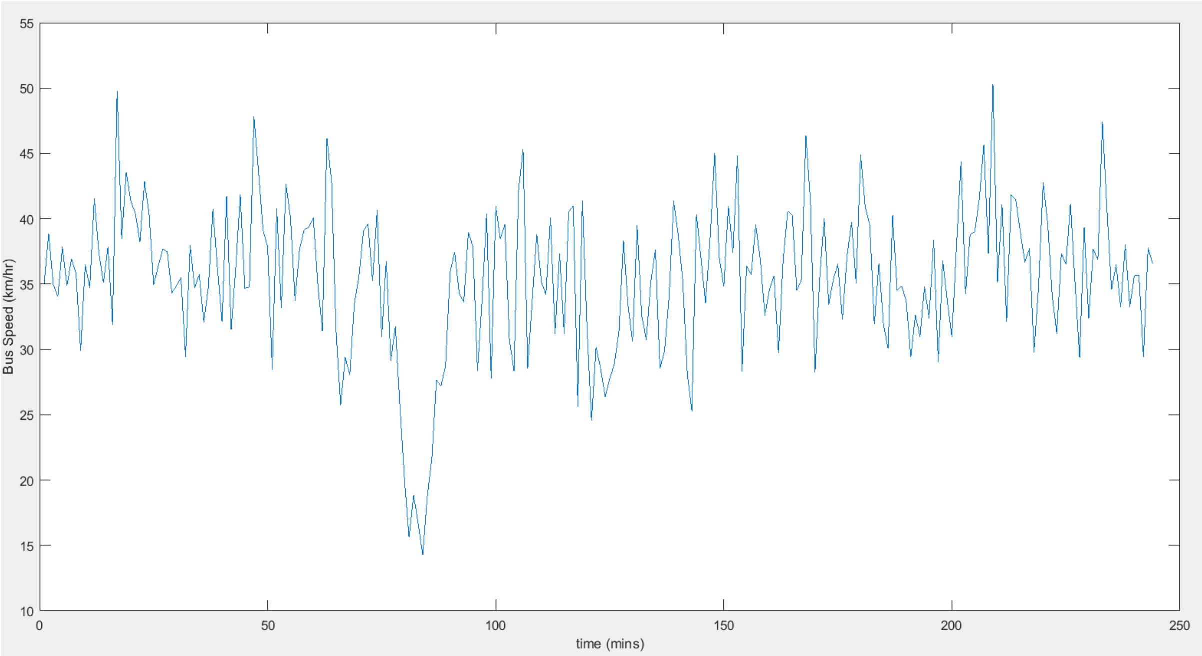
a) *CardID*: unique ID of the smart card (Go Card),
b) *TimeOn*: time the card is touched on (boarded)
c) *TimeOff*: Time the card is touched off (alighted)
d) *BusRouteID*: Route number of the bus
e) *BoardingStop:* Stop ID of the stop from where the passenger has boarded
f) *AlightingStop*: Stop ID of the stop from where the passenger has alighted
g) *BoardingTime*: Time the card is touched on
h) *AlightingTime*: Time the card if touched off
i) *ScheduleStartTime*: Schedule start time of the bus

The above raw data is processed and travel time of the bus between two stops is estimated based on the first *touch-on/touch-off* between the two stops. For details refer to Kieu et al., (2015). We generate bus speed time series with a bin of 5 minutes.

5.2.1 Ground Truth: Bus Time Series under high frequency buses

As mentioned above, the study site has high frequency bus routes. The time series of bus speed estimated using the entire data set contains bus speed with only 2.8% missing values. For the current analysis, we assume that this time series is the ground truth that should be used to compare with the low frequency buses for bus speed estimation and prediction. **Figure *3*** illustrate sample ground truth of the bus speed.

Figure 3: Sample bus speed from Go card



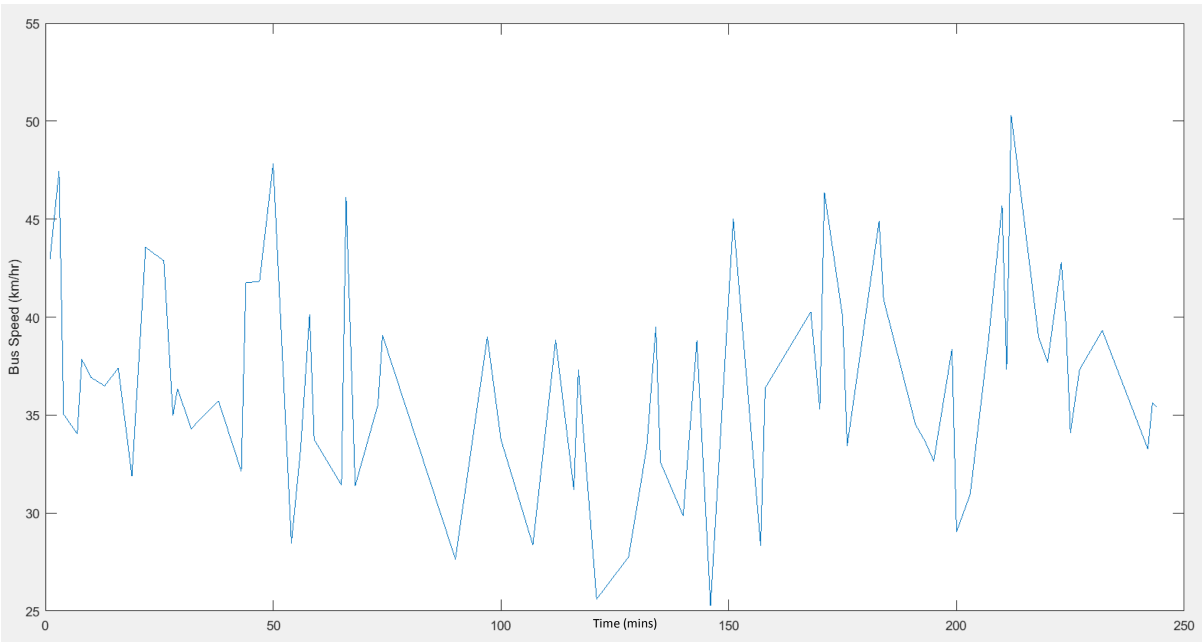
5.2.2 ***Available bus speed: Bus Time Series under low frequency of buses***

To represent a case with low frequency of buses, we only consider one route to generate another time series of bus speed. This has significant number of missing values. This series will be used for model development. For the available bus speed, we will fill the missing values using a) Case-1: Linear interpolation; and b) Case-2: knowledge from car speed.

5.2.2.1 Case-1 (LI): Estimating missing values using Linear Interpolation

An example of bus speed missing values filled with linear interpolation for June 2015 is shown in **Figure 4**.

Figure 4: Sample time series of the bus speed using linear interpolation



5.2.2.2 *Case-2 (CI): Estimating missing values using knowledge from car speed*

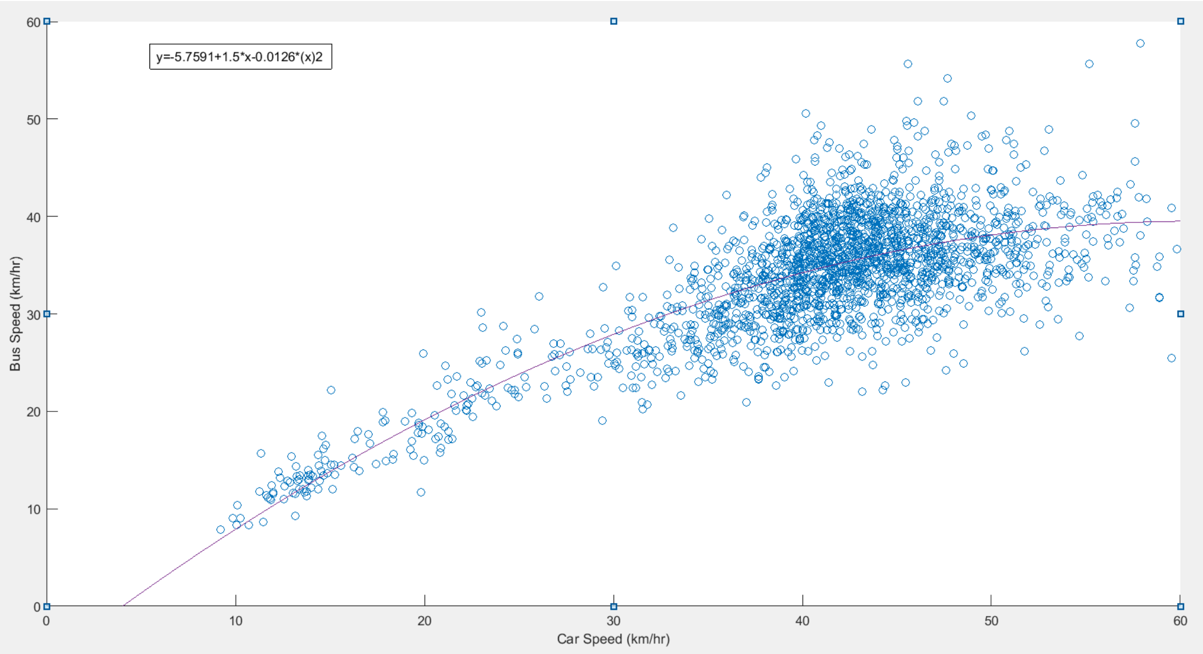
Here we need to first estimate the relationship between the bus and car speed. For this, data is extracted during the periods when both bus and car speed data points are available. The figure below represents the extracted bus and car speed data from the study site. For this paper, we develop following regression model to relate the bus speed with car speed, which can be used to demonstrate the benefit of integration of car speed for bus speed estimation and prediction.

$$V_{bus}=-5.7591+1.5*V_{car}-0.0126*(V_{car})^2 \quad (4)$$

Where, V_{bus}= Bus Speed

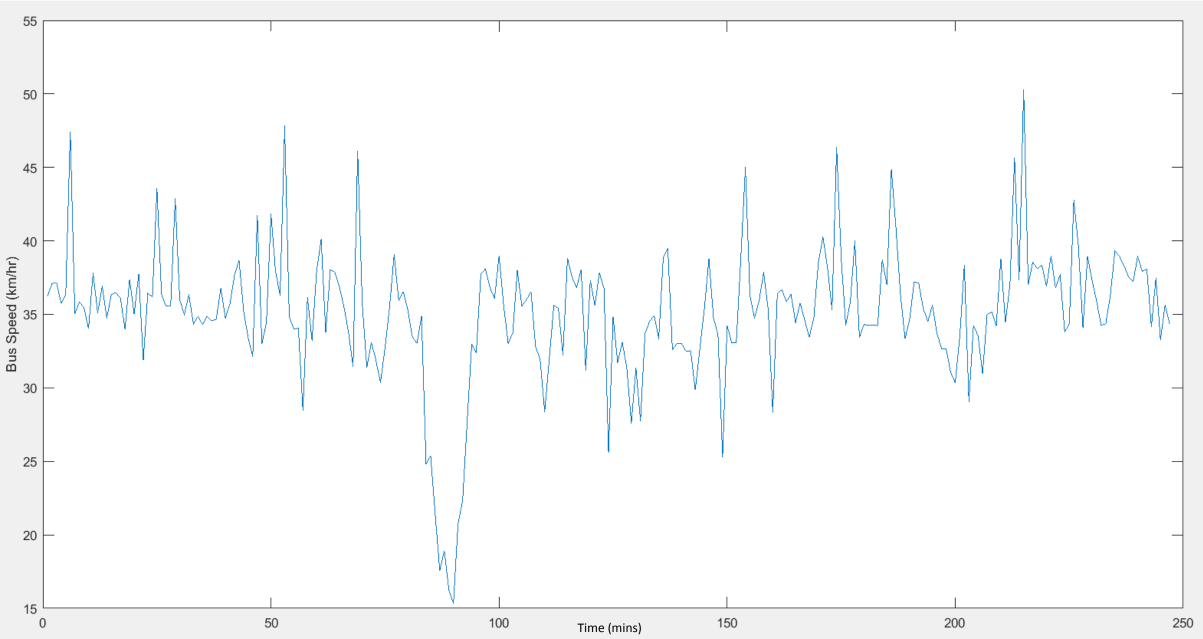
V_{car}= Car Speed

Figure 5: Bus car speed relationship



The above developed regression model is used to fill the missing values in bus time series with an estimate from the car speed. An example series is provided in **Figure 6**.

Figure 6: Sample time series of the bus speed integrating car speed



6 Results

6.1 Estimation results

Here we compare the ground truth with results from Case-1 (LI) and Case-2 (CI). Only data points where the estimate of the bus speed is used by either of the two cases is used for comparison.

The proposed models are validated with the validation data that has not been used for model development. The performance is evaluated in terms of Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) defined in terms of actual speed (V_A) and estimated speed (V_E).

$$RMSE=\sqrt{\frac{\sum_{i=1\ to\ N}(V_{A,i}-V_{E,i})^2}{N}} \quad (5)$$

$$MAPE=\frac{\sum_{i=1\ to\ N}\left(\frac{|V_{A,i}-V_{E,i}|}{V_{A,i}}\right)}{N} \quad (6)$$

Where N is the number of data points used for evaluation.

The following table presents the performance of Case 1 and Case 2.

Table 1: performance of Case 1 and Case 2

	Case 1- Linear interpolation	**Case 2- Integrating with car speed**
MAPE (%)	**9.7**	**6.9**
RMSE (Km/hr)	**4.9**	**3.6**

From the results, it is clearly seen that estimation results are better for case 2 than case 1 in terms of MAPE and RMSE. The results are plotted in **Figure *7*** and **Figure *8*** where x axis represents actual speed in km/hr and y axis represents estimated speed in km/hr.

Figure 7: Actual vs estimated speed for case 1

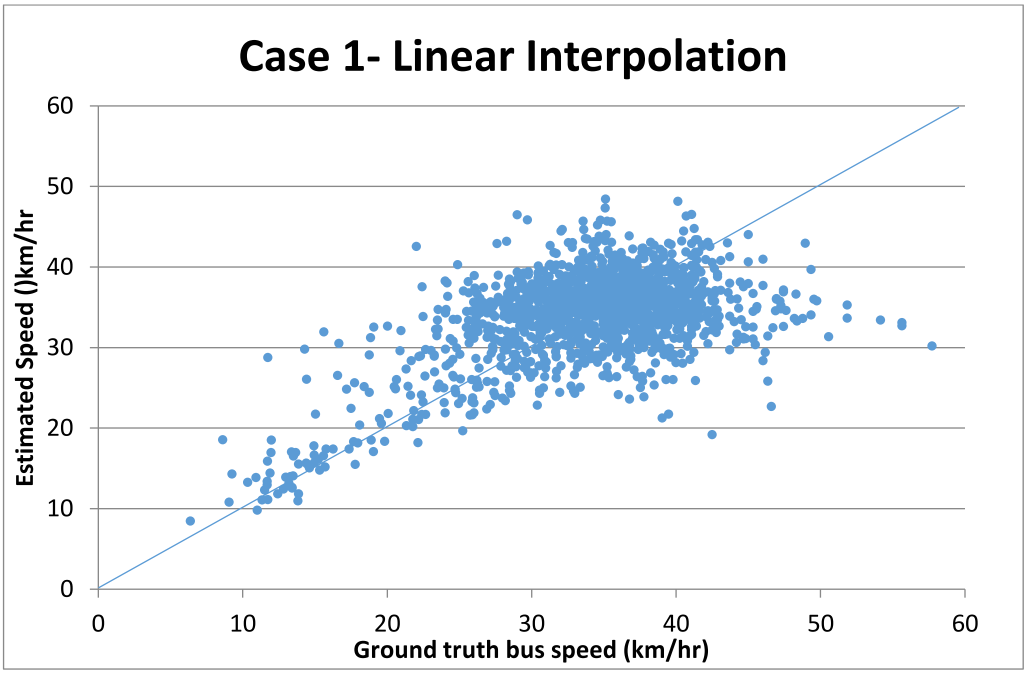
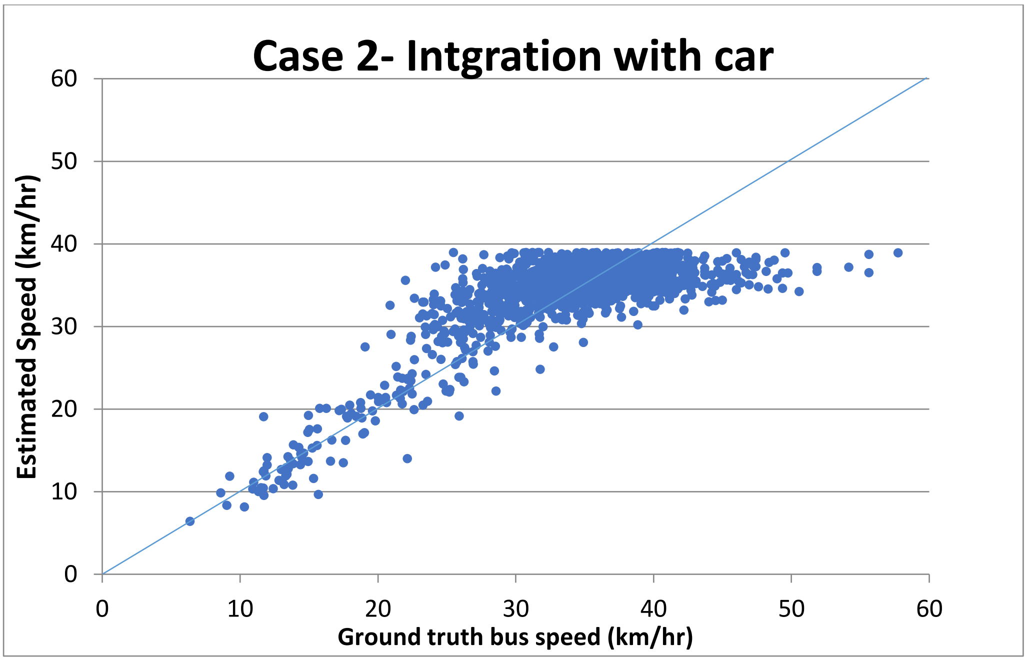


Figure 8: Actual vs estimated speed for case 2



6.2 Prediction model- ANN

Here, the objective is to check the performance of bus speed time series in terms of the accuracy of short term speed prediction. For this we chose Artificial Neural Network (ANN) to evaluate the accuracy of bus speed time series. The reasons for selecting ANN are given below:

- ANNs are well-suited for problems whose solutions require knowledge that is difficult to specify.
- ANNs can generalize.

In this study, Non Linear Autoregressive Neural Network (NAR) has been used. These networks is that they can accept dynamic inputs represented by time series sets. Time series forecasting using neural network (NN) is a non-parametric method, which means that knowledge of the process that generates the time series is not indispensable. The NAR model uses the past values of the time series to predict future values.

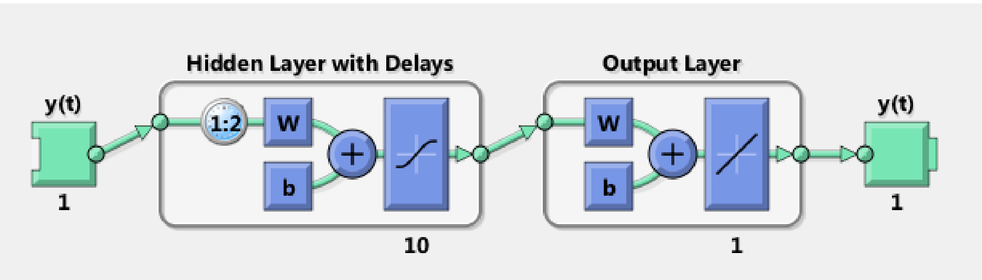
6.2.1 *NAR Model*

In the majority of cases, time series applications are characterized by high variations and fleeting transient periods. This fact makes it difficult to model time series using a liner model, therefore a nonlinear approach should be suggested. A nonlinear autoregressive neural network applied to time series forecasting, describes a discrete, non-linear, autoregressive model that can be written as follows:

$$y(t) = f\big(y(t-1), y(t-2), \dots \dots \dots y(t-d)\big) + \varepsilon(t) \qquad (7)$$

This formula describes how a NAR network is used to predict the value of a data series y at time t, y (t), using the d past values of the series. The function f is unknown in advance, and the training of the neural network aims to approximate the function by means of the optimization of the network weights and neuron bias. Finally, the term ε (t) stands for the error of the approximation of the series y at time t. The topology of a NAR network is shown in **Figure *9***.

Figure 9: Topology of NAR Network



The d features y (t−1), y (t−2) . . . y (t−d), are called feedback delays. The number of hidden layers and neurons per layer are completely flexible, and are optimized through a trial-and-error procedure to obtain the network topology that can provide the best performance.

For training, Bayesian Regularization procedure is used. This algorithm typically requires more time, but can result in good generalization for difficult, small or noisy datasets. Training stops according to adaptive weight minimization (regularization). Specifications of ANN model are presented in a table below.

Table 2: Specifications of ANN model

Type of ANN	No of Hidden Neuron	d	Training Function
NARNN	10	2	Bayesian Regularization

6.3 Prediction modelling results

For prediction, we develop the model using the available bus speed and validate the model using the validation data from the ground truth. ANN is selected as the tool to test the prediction performance. Both models are tested for different prediction horizon ranging from t+5, t+10, t+15, t+20, t+25, t+30 minutes.

6.3.1 *Case-1- Estimation using Linear Interpolation: Prediction results*

In this case model is developed on the data where bus speed missing values are estimated with linear interpolation and the developed model is model 1. Model 1 is used to predict future travel speed for different time horizon and the results are presented in **Table *3***.

Table 3: Prediction results for Model 1

Prediction horizon (mins)	MAPE (%)	RMSE(km/hr)
t+5	15	6.3
t+10	16.8	6.7
t+15	19.9	7.4
t+20	21.1	7.9
t+25	23.1	8.4
t+30	24.4	8.7

6.3.2 *Case-2- Estimation using an estimate from car speed: Prediction results*

In this case model is developed on the data where bus speed missing values are estimated with the knowledge from car speed and the developed model is model 2. Model 2 is used to predict future travel speed for different time horizon and the results are presented in **Table *4***.

Table 4: Prediction results for Model 2

Prediction horizon (mins)	MAPE (%)	RMSE(km/hr)
t+5	13	5.3
t+10	15.8	6.0
t+15	17.7	6.5
t+20	19.3	7.1
t+25	20.9	7.6
t+30	22.5	7.9

From both model 1 and model 2, it can be seen that the performance is decreasing when the lead time increases for both mean absolute percentage error and root mean square error. In every case for instance for lead time 5 mins to 30 mins model 2 (integrated car speed model) performs better than model 1(linear interpolation). Performance of model 1 and model 2 is shown in following figures:

Figure 10: Comparison of Model 1and Model 2 in terms of MAPE

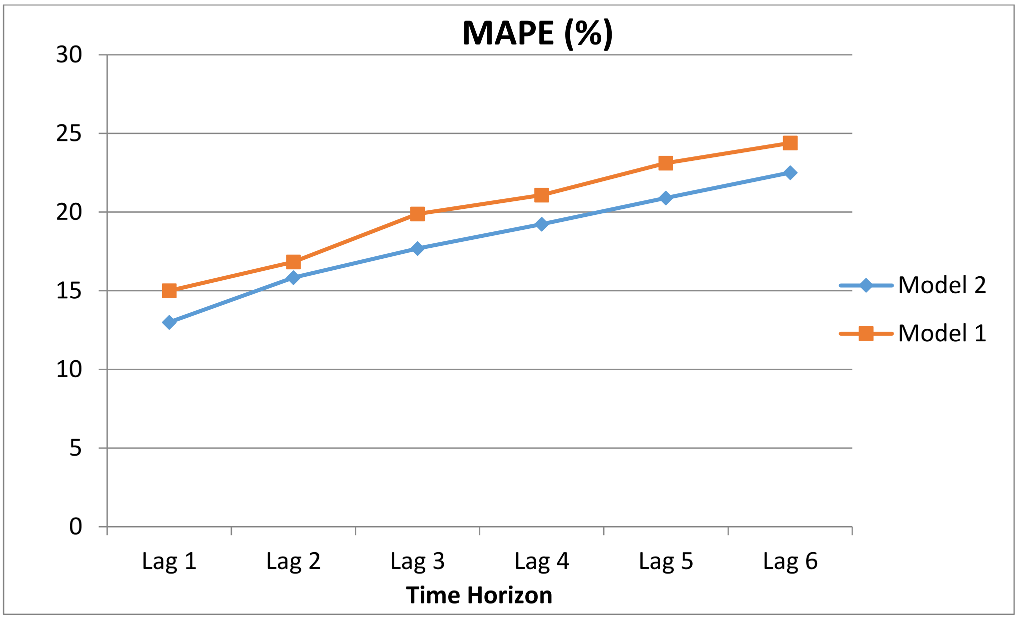
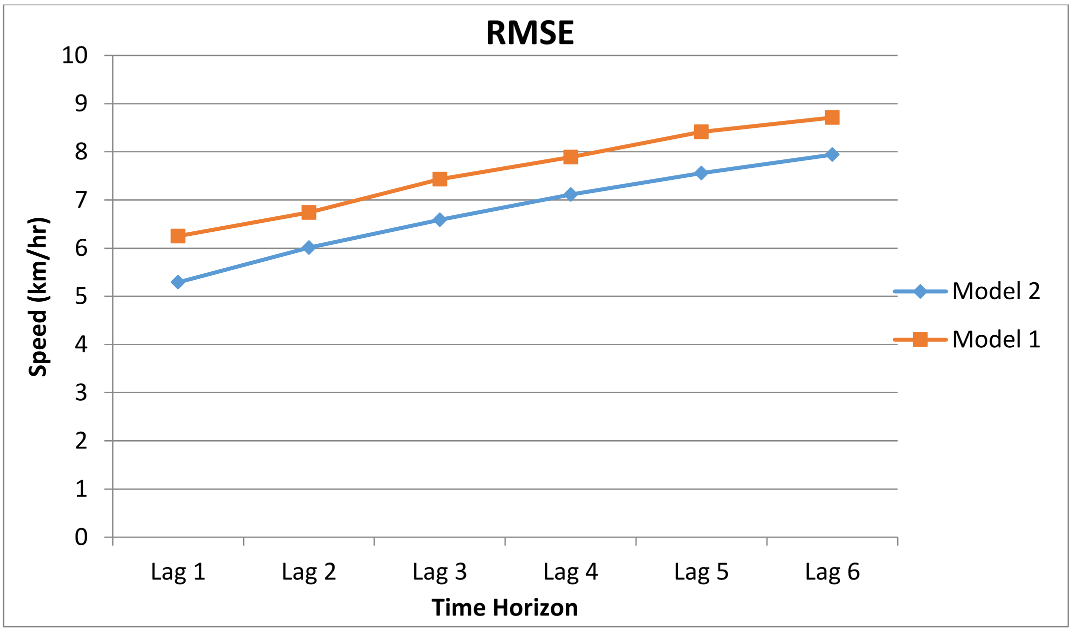


Figure 11: Comparison of Model 1and Model 2 in terms of RMSE



7 Conclusion

Understanding the relationship between multi-modals of transport could benefit strategic traffic planners and road operators. In this paper, the real data (Bluetooth and Go card) from Brisbane, is exploited to first empirically model the relationship between bus and car speed and thereafter utilised to develop database for bus time series estimates along urban corridor. The database is finally applied for short term bus speed (travel time) prediction on urban corridor. The results indicate an improvement of 3% for bus speed estimation and more than 2 % for bus speed prediction for different time horizon.

In future studies, other variables which may contribute to the difference between bus and car speeds such as the number of boarding/alighting passengers, type of bus, etc. will be explored.

Reference

BAE, S. 1995. Dynamic Estimation of Travel Time on Arterial roads by Using Automatic Vehicle Location (AVL) Bus as a Vehicle Probe.

BERTINI, R.L., and TANTIYANUGULCHAI, S. 2004. "Transit buses as traffic probes: Use of geolocation data for empirical evaluation." *Transportation Research Record: Journal of the Transportation Research Board* no. 1870 (-1):35-45.

BHASKAR, A. & CHUNG, E. 2013a. Fundamental understanding on the use of Bluetooth scanner as a complementary transport data. *Transportation Research Part C: Emerging Technologies,* 37**,** 42-72.

BHASKAR, A., CHUNG, E. & DUMONT, A.-G. 2010. Analysis for the Use of Cumulative Plots for Travel Time Estimation on Signalized Network. *International Journal of Intelligent Transportation Systems Research,* 8**,** 151-163.

BHASKAR, A., CHUNG, E. & DUMONT, A. G. Arterial travel time estimation: Revisiting the classical procedure. Australasian Transport Research Forum, 2011 Adelaide, South Australia.

BHASKAR, A., QU, M. & CHUNG, E. 2014a. A Hybrid Model for Motorway Travel Time Estimation- Considering Increased Detector Spacing. *Transportation Research Record: Journal of the Transportation Research Board,* 2442.

BHASKAR, A., QU, M. & CHUNG, E. 2015. Bluetooth Vehicle Trajectory by Fusing Bluetooth and Loops: Motorway Travel Time Statistics. *IEEE Transactions on Intelligent Transportation Systems,* 16**,** 113-122.

BHASKAR, A., TSUBOTA, T., KIEU, L. M. & CHUNG, E. 2014b. Urban traffic state estimation: Fusing point and zone based data. *Transportation Research Part C: Emerging Technologies,* 48**,** 120-142.

CATHEY, F. W. & DAILEY, D. J. 2001. Transit Vehicles as Traffic Probe Sensors. *IEEE Intelligent Transportation Systems Conference Proceedings.* Oakland (CA), USA.

CHIEN, S. I. & KUCHIPUDI, C. M. 2003. Dynamic Travel Time Prediction with Real-Time and Historic Data. *Journal of transportation engineering,* 129**,** 608-616.

CHIEN, S. I., DING, Y. & WEI, C. 2002. Dynamic Bus Arrival Time Prediction with Artificial Neural Networks. *Journal of transportation engineering,* 128**,** 429-438.

FADAEI, M., CATS, O. & BHASKAR, A. 2017. A hybrid scheme for real-time prediction of bus trajectories. *Journal of Advanced Transportation***,** n/a-n/a.

FEI, X., LU, C.-C. & LIU, K. 2011. A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction. *Transportation Research Part C: Emerging Technologies,* 19**,** 1306-1318.

JEONG, R. & RILETT, L. R. 2004. Bus Arrival Time Prediction Using Artificial Neural Network Model. *IEEE Intelligent Transporlatlon Systems Conference.* Washington, D.C., USA.

KHOEI, A. M., BHASKAR, A. & CHUNG, E. Travel time prediction on signalised urban arterials by applying SARIMA modelling on Bluetooth data. 36th Australasian Transport Research Forum (ATRF), 2 - 4 October 2013 Brisbane, Australia.

KIEU, L. M., BHASKAR, A. & CHUNG, E. 2012. Bus and car travel time on urban networks: integrating bluetooth and bus vehicle identification data. *25th ARRB Conference – Shaping the future: Linking policy, research and outcomes.* Perth, Australia.

KIEU, L. M., BHASKAR, A. & CHUNG, E. 2015. Empirical modelling of the relationship between bus and car speeds on signalised urban networks. *Transportation Planning and Technology,* 38**,** 465-482.

LEVINSON, H. 1983. "Analyzing transit travel time performance." *Transportation*

Research Record (915).
LIN, Y., YANG, X., ZOU, N. & JIA, L. 2013. Real-Time Bus Arrival Time Prediction: Case Study for Jinan, China. *Journal of Transportation Engineering,* 139**,** 1133-1140.
MAZLOUMI, E., ROSE, G., CURRIE, G. & SARVI, M. 2011. An Integrated Framework to Predict Bus Travel Time and Its Variability Using Traffic Flow Data. *Journal of Intelligent Transportation Systems,* 15**,** 75-90.
MCKNIGHT, C. E., LEVINSON, H. S., OZBAY, K., KAMGA, C. & PAASWELL, R. E. 2004. The Impact of Traffic Congestion on Bus Travel Time in Northern New Jersey. *Transportation Research Record: Journal of the Transportation Research Board,* 1884**,** 27-35.
TSUBOTA, T., BHASKAR, A., CHUNG, E. & BILLOT, R. 2011. Arterial traffic congestion analysis using Bluetooth duration data. *34th Australasian Transport Research Forum (ATRF).* Adelaide, South Australia, Australia.

VAN HINSBERGEN, C. P. I., VAN LINT, J. W. C. & VAN ZUYLEN, H. J. 2009. Bayesian committee of neural networks to predict travel times with confidence intervals. *Transportation Research Part C: Emerging Technologies,* 17**,** 498-509.