Transport
Volume 167 Issue TR3

Pattern identification based bus arrival time prediction
Kumar and Vanajakshi

ice | proceedings

Proceedings of the Institution of Civil Engineers
Transport 167 June 2014 Issue TR3
Pages 194–203 http://dx.doi.org/10.1680/tran.12.00001
Paper 1200001
Received 29/12/2011 Accepted 20/08/2012
Published online 10/04/2013
Keywords: developing countries/statistical analysis/transport management

ice
Institution of Civil Engineers
publishing

Pattern identification based bus arrival time prediction

1 **Selvaraj Vasantha Kumar** ME
PhD Research Scholar, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India

2 **Lelitha Vanajakshi** MTech, PhD
Assistant Professor, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India


1


2

The provision of accurate bus arrival time information is one of the major components in advanced public transportation systems, which many of the metropolitan cities in developing countries are trying to implement to increase the public transit usage. The effectiveness of such a system largely depends on the reliability of the information provided to the public. For such reliable information to be generated, the prediction technique used should be able to make accurate predictions, which in turn depends on the input data used for prediction. The present study is an attempt to explore these two areas, namely the identification of suitable input data by analysing trip-wise, daily and weekly patterns of bus travel times through valid statistical tests and the development of an accurate bus arrival prediction model using a popular time series technique called exponential smoothing. The performance evaluation using 90 actual bus trip data shows that the use of suitable input data into the prediction model yields better results with mean absolute percentage error of 12, and for 77% of the time the deviation of predicted arrival time with respect to actual arrival time is within the user acceptable range of ± 5 min.

Notation

b_{obs}	observed arrival time
b_{pre}	predicted arrival time
$\bar{d}$	mean of the differences
MAPE_{bs}	mean absolute percentage error between observed and predicted arrival time at bus stop bs
n	sample size
s_{d}	sample standard deviation
S_t	forecast at time t
S_{t-1}	forecast at time $t-1$
$\hat{x}_k$	predicted travel time of the test vehicle (TV) in the kth section
x_{k-1}	observed travel time in the previous $(k-1)$th section
$\hat{x}_{k-1}$	predicted travel time of the TV in the $(k-1)$th section
y_{t-1}	observation at time $t-1$
z	test statistic
α	smoothing constant

1. Introduction

Informing public transit users about the expected arrival time of the next bus is one of the major components of advanced public transportation systems (APTS), which are expected to attract more travellers onto public transit, thereby reducing congestion on urban roads. However, for this to be successful, the information on expected bus arrival time provided should be reliable, otherwise customers will reject the system (Schweiger, 2003). The reliability of the information being provided depends on the prediction technique used, which in turn depends on the input data used for prediction. The present study is an attempt to explore these two areas, namely the identification of suitable input data by analysing trip-wise, daily and weekly patterns of bus travel time through valid statistical tests and the development of a simple and accurate bus arrival prediction model using a popular time series technique called exponential smoothing.

The input data for the bus arrival time prediction model can be generally classified as traffic-based or transit-based. The traffic-based data include the current traffic conditions (e.g. intersection delays, weather conditions etc.) or historical traffic conditions (e.g. traffic state by day and time of day in the past). The transit-based data include real-time operating data from the last several buses on a particular route of interest (e.g. running time between stops) or historical bus operations data (e.g. running times between bus stops by day and time of day). In order to incorporate the effect of traffic-based (either current or historic) characteristics into the transit-based data, three kinds of patterns, namely the trip-wise, daily and weekly patterns of bus travel time, were proposed and analysed in the present study. A brief description of each of these patterns is explained in the following paragraph.

In the trip-wise pattern analyses the assumption is that the current trip has a similar pattern as many previous trips of the same day. This can capture traffic conditions that are specific to the day such as incidents that happened in that day and temporary diversions. The daily pattern checks the significance of many same-time trips on previous days for the current trip. The weekly pattern checks whether the same-day/same-time trip of previous week(s) has a similar pattern as the current trip. These can capture repeating patterns that are dependent on the time of trip/day of the trip, such as peak period congestion, Monday morning congestion and so on. The present study reports a systematic statistical analysis of these patterns in order to identify the most suitable trips which can be used as an input into the prediction model.

The studies on bus arrival time prediction can use any of the following techniques: historic and real-time approaches (Chien and Kuchipudi, 2003); machine learning techniques such as artificial neural networks (ANN) (Chien *et al.*, 2002; Jeong and Rilett, 2004) and support vector machines (SVM) (Bin *et al.*, 2006); model-based approaches using Kalman filtering (Dailey, 1999; Vanajakshi *et al.*, 2009); and statistical methods such as regression analysis (Bo *et al.*, 2009; Jeong and Rilett, 2005; Patnaik *et al.*, 2004; Ramakrishna *et al.*, 2006) and time series techniques (Bhandari, 2005; Suwardo *et al.*, 2010).

Many of the above techniques such as regression, SVM, ANN and time series analysis carry out the prediction based on the patterns existing in the data and hence require a sound database. The studies which are based on these techniques usually classify the input data by different time periods of the day, and use the pattern while predicting a particular trip on a day. Although it takes into account the traffic conditions specific to that time period of the day, the question of whether the bus trip of months before is significant in predicting a current trip was not addressed. The prediction models which are based on filtering techniques need fewer data for model development, to the level of data from the previous one or two trips. It is to be noted here that, in all of the studies listed above, the input data for the prediction model were arbitrarily selected without any statistical basis. However, identifying the most suitable trips in predicting the next bus travel time and using them in the prediction models will definitely improve the prediction accuracy. The present study reports research in this area to analyse bus travel time patterns by performing valid statistical tests in order to identify the most suitable trips which can be used as an input into the prediction model for predicting the next bus travel time.

The time series approach of bus arrival time prediction mainly used Box-Jenkins models (Bhandari, 2005; Suwardo *et al.*, 2010) such as autoregressive (AR), moving average (MA) or a combination of AR and MA models (Arma, Arima etc.). The use of classical time series methods such as exponential smoothing was not attempted for bus arrival time prediction. The major disadvantage of using Box-Jenkins models is that without time series modelling software, parameter estimation and forecasting is computationally prohibitive (Nickerson and Madsen, 2005). Thus, the necessity of dependence on time series modelling software (e.g. R, ITSM, SAS) restricts the applicability of Box-Jenkins models for real-time applications like bus travel time/arrival time prediction. Also, the family of Box-Jenkins models may be difficult to explain to others (Binoy, 2011). Due to these disadvantages associated with Box-Jenkins models, classical time series methods such as exponential smoothing techniques are still being widely used in many areas such as electrical engineering (Temraz *et al.*, 1996; Damrong and Churueang, 2005), environmental and social sciences (Booth and Tickle, 2008; Mondal *et al.*, 2010; Wexler, 1996), ecology (Shurin, 2010), economics (Arsham and Shaw, 1985) and manufacturing (Gardner and Saiz, 2002). These smoothing methods were developed in the 1960s before the Box-Jenkins models were available and are still the most popular forecasting methods used in business and industry. The major advantage of using exponential smoothing when compared to ARIMA models is that they are easy to understand and apply and with a simple spreadsheet program (or programming in Matlab or C, C++), the method can easily be implemented eliminating the requirement for time series modelling software. Hence, the present study tries to explore the usefulness of the exponential smoothing method for the problem of bus arrival time prediction using the most suitable bus trips identified by analysis of travel time patterns of public transit buses under heterogeneous traffic conditions for real-time implementation.

The following section gives an overview of the data collection and extraction part. Section 3 explains the methodology for travel time pattern analysis and its results. The equations of the prediction model are explained in Section 4 along with validation of the prediction model using real-world data, followed by concluding remarks in Section 5.

2. Data collection and extraction

The study stretch selected for the present study was route number 5C, which connects the Parrys bus depot in the northern part of Chennai, and the Taramani bus depot in the southern part of Chennai city. The time headway between the buses was 30 min during most of the day. The total route length is 15 km and the approximate travel time to cover the total stretch is 70 min during peak hours and 50 min during off-peak hours. There are 21 bus stops and 14 signalised intersections on this route. The total route comprises road links of different categories with varying volume levels and numbers of lanes.

The data collection involved the automatic vehicle location (AVL) of seven public transit buses reporting every 5 s from 8 a.m. to 8 p.m. using permanently fixed global positioning system (GPS) units over a period of three months from September to November 2010. Fifteen days from 15 November 2010 to

29 November 2010 were taken as the period for the analysis. Since the time headway of buses is 30 min, approximately one trip for every 30 min was available from 8 a.m. to 8 p.m. Thus, the data consisted of a total of 360 trips (24 trips/day × 15 days). Each trip (to be called 'output trip' below) was compared with 10 preceding trips of the same day, the previous 10 days' trips with the same starting time as that of the output trip and the previous three weeks' trips of the same-day/same-time to identify the trip-wise, daily and weekly travel time patterns. Each of these preceding trips will be called 'input' trips below. The data extraction involved extracting travel times for each 100 m section for the above trips in order to facilitate pair-wise comparison while applying statistical tests.

For validating the prediction model developed using the exponential smoothing technique, a sample of 105 trips spanned across 5 days during the months of September to November 2010 was considered. For predicting the travel time/arrival time at bus stops for each of these 105 trips, the corresponding selected inputs, namely the previous week(s) same time trip or previous day(s) same time trip or previous trip(s) of the same day, are used. The validation of the proposed model involved the comparison of observed arrival time of the bus with the predicted arrival time at 21 bus stops along the route, for each of the 105 trips. Hence, the data extraction also involved extracting each 100 m section travel time along the study stretch and arrival time at 21 bus stops for all the 105 trips of the five days considered.

3. Travel time pattern analysis

The purpose of bus travel time pattern analysis is to find out whether many previous trip(s) of the same day or many previous days' same-time trips or many previous weeks' same-day/same-time trips are significant in predicting the next bus travel time. For this, the statistical test, namely the 'z-test for the mean of a population of differences for "paired" samples data' was used for hypothesis testing at a 5% significance level. The 'paired' sample means, comparing each 100 m section travel time of output trip to that of input trip to check whether the mean of the population of differences is significantly different from zero or not. Thus the hypothesis testing results in whether the $(t-n)$ previous trip(s) of the same day or $(t-d)$ previous day(s) same-time trip or $(t-w)$ previous week(s) same-day/same-time trip is significant in predicting the output trip t, where n and d vary from 1 to 10 for the trip-wise pattern and daily pattern respectively and w varies from 1 to 3 for the weekly pattern. The null hypothesis was that the mean of the population of differences is not significantly different from zero. In other words, the mean of the differences of each 100 m section travel time of output trip and input trip is zero, that is, the $(t-n)$ or $(t-d)$ or $(t-w)$ trip is significantly similar to the output trip t. The alternative hypothesis is that differences are not zero. The test statistic is given by

1. $$z = \bar{d}/\frac{s_\mathrm{d}}{\sqrt{n}}$$

where $\bar{d}$ is the mean of the differences, s_d is the sample standard deviation and n is the sample size. At 5% significance level, if the calculated z is within the range of $-1{\cdot}96$ and $+1{\cdot}96$, the mean of the population of differences can be considered as not significantly different from zero. For analysing the trip-wise pattern, 3600 z-values were calculated (24 trips/day × 10 previous trips × 15 days (from 15–29 November 2010)) using Equation 1. Similarly 3600 z-values were calculated for the daily pattern, and 1080 z-values were calculated for the weekly pattern. Thus, a total of 8280 z-values were calculated using Equation 1. In order to derive some conclusions from the large collection of 8280 z-values, the percentage of the number of times the null hypothesis is accepted is calculated for each $(t-n)$, $(t-d)$ and $(t-w)$ case. If the percentage is higher, it implies that the input trip ($(t-n)$ or $(t-d)$ or $(t-w)$) is more significant in predicting the output trip t. From the percentage values, we can also say whether trip-wise, daily or weekly patterns exist. Thus, by looking at the percentage values for each $(t-n)$, $(t-d)$ and $(t-w)$ case, one can identify which are the most important trips that need to be considered as an input for predicting the next bus travel time in the bus arrival time prediction model.

The results for trip-wise patterns are shown in Figure 1, which shows that the immediate previous bus $(t-1)$ has the highest percentage acceptance of 62·33, when compared to the percentage acceptance of previous 10 trips of the same day. This clearly shows that the preceding bus is more significant in predicting the next bus travel time. Although the effect of the $(t-2)$ bus is less than that of the $(t-3)$, overall the previous three buses $(t-1)$, $(t-2)$ and $(t-3)$ can be considered to have the highest percentage acceptance values when compared to other percentage acceptance values in the trip-wise pattern. It is interesting to see that the percentages gradually decrease from $(t-3)$ to $(t-6)$ and then start increasing again. This clearly shows the repeating nature of peak, off-peak cycles in a particular day.

The percentage acceptances for the daily pattern are also shown in Figure 1. It can be seen that the percentage for the $(t-7)$ day is higher when compared to all other days, also that the $(t-7)$ day is the same as the $(t-1)$ week thus showing the same percentage acceptance values. This clearly indicates that there exists a strong weekly pattern, with the previous week same-day/same-time trip being more significant in predicting the current trip than any of the previous day's same time trips. Similarly for the weekly pattern, the percentage acceptance for $(t-1)$ week, $(t-2)$ week, $(t-3)$ week are 69·23, 66·22 and 58·22 respectively as shown in Figure 1, with $(t-1)$ week and $(t-2)$ week being exceptionally higher than for $(t-3)$ week.

Based on the pattern analysis results as described above, one can find out the significant or most important trips that could be considered as inputs for the bus arrival time prediction model. Based on the trip-wise pattern results, the three previous buses were considered as inputs as they exhibit the highest percentage acceptance values of 62·33, 54·71 and 60·99 respectively. Among

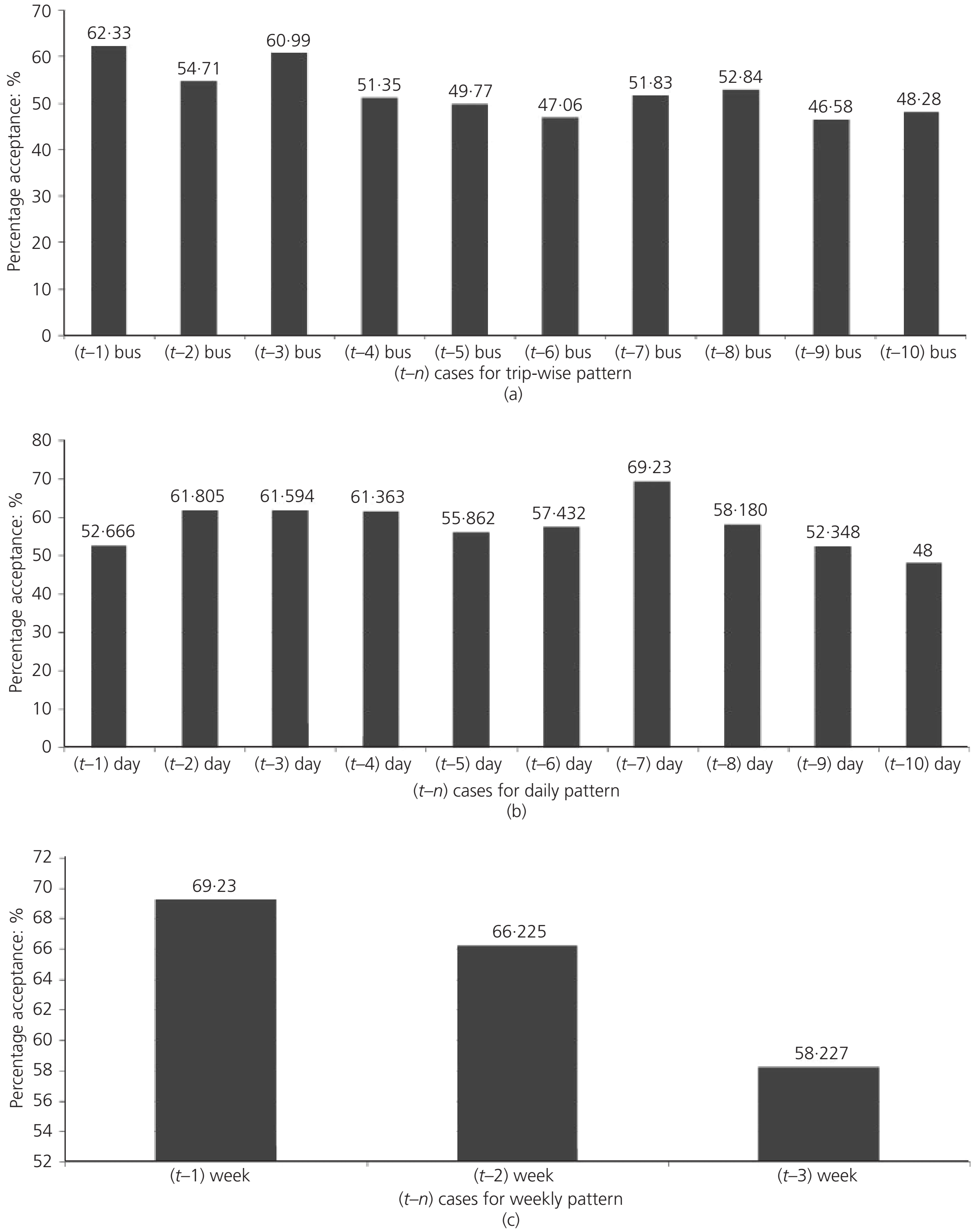

Figure 1. Percentage acceptance for: (a) trip-wise, (b) daily and (c) weekly patterns

the daily and weekly pattern, it is clear from Figure 1 that the weekly pattern dominates the daily pattern with $(t-1)$ week and $(t-2)$ weeks having the highest percentage acceptance values. Thus, the previous two weeks' same-day/same-time trips (called W1 and W2 below) and the previous three trips of the same day (called PV1, PV2 and PV3 below) were identified to be the significant or most important trips and were included as the inputs in the prediction model developed using the simple exponential smoothing technique as explained in the following section.

4. Bus arrival time prediction model

The model for predicting the bus arrival time was developed using one of the classical time series techniques called simple exponential smoothing. The exponential smoothing weights past observations with exponentially decreasing weights to forecast future values. In other words, recent observations are given relatively more weight in forecasting than the older observations. The basic equation of simple exponential smoothing is given as

$$S_t = \alpha y_{t-1} + (1-\alpha)S_{t-1} \tag{2}$$

where S_t is the forecast at time period t, y_{t-1} is the observed value at time $(t-1)$ and S_{t-1} is the forecast value at time $(t-1)$. The parameter α is the smoothing constant $(0 \leqslant \alpha \leqslant 1)$.

The study stretch of 15 km was divided into 100 m sections of uniform length to facilitate the travel time prediction in each of these sections using the exponential smoothing as explained below. For this problem, Equation 2 can be rewritten as

$$\hat{x}_k = \alpha\, x_{k-1} + (1-\alpha)\hat{x}_{k-1} \tag{3}$$

where $\hat{x}_k$ is the predicted travel time of the test vehicle (TV) in the kth section, x_{k-1} is the observed travel time in the previous $(k-1)$th section (the weighted average of the previous two weeks same-day/same-time trips (W1, W2), and the previous three trips (PV1, PV2, PV3) of the same-day travel time in $(k-1)$th section), $\hat{x}_{k-1}$ is the predicted travel time of the TV in the $(k-1)$th section. Since the predicted TV travel time in the first 100 m section is unknown, the observed TV travel time is taken in its place. In order to find an optimum α-value as well as the optimum relative weights to be used for the inputs W1, W2 and PV1, PV2, PV3, the following procedure was adopted.

Out of 105 trips spanning five days, 15 trips of one sample day (16 November) were used for finding out the optimum weights and α-value and the remaining 90 trips in four days were kept for validation. A total of nine cases were considered with each case representing the relative weights given to W1, W2 average section travel time and PV1, PV2, PV3 average section travel time. For example, for the first case, the weight assigned is 0·1 for the average of W1, W2 and 0·9 for the average of PV1, PV2, PV3. The weight for W1, W2 average was increased by an increment of 0·1 and the weight for PV1, PV2, PV3 average was decreased by 0·1 for the subsequent cases. Under each case, nine sub-cases were considered with varying α-values ranging between 0·1 and 0·9. The exponential smoothing model as shown in Equation 3 was executed for all the 15 trips of 16 November using the 81 cases (9×9) with the relative weights for the inputs and varying α-values. The mean absolute percentage error (MAPE) is used as a measure of estimation accuracy and is calculated using

$$\text{MAPE}_{\text{bs}} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{b_{\text{pre}} - b_{\text{obs}}}{b_{\text{obs}}}\right| \times 100 \tag{4}$$

where MAPE_{bs} is the mean absolute percentage error between observed and predicted arrival time at various bus stops (bs ranges from bus stop 1 to 21). The term b_{pre} and b_{obs} are the predicted and observed arrival times of the bus for the bus stop bs, and n is the number of trips in the given day.

The results of the MAPE between observed and predicted arrival times are shown in Figure 2, where the MAPE gradually decreases with increasing weights for W1, W2 average section travel times and a saturation stage is reached when the weight assigned is 0·8. This clearly shows that the use of the previous two weeks' same time trips data has more impact for the next bus arrival time prediction than the previous three buses of the same day. This was even reflected during pattern analysis, with $(t-1)$ week and $(t-2)$ week having the highest percentage acceptance values, when compared to others. Taking this as the best option, the analysis to choose the best α-value was carried out. It can be observed that the MAPE gradually decreased with increasing α-values, with the reduction in error being very high initially and reaching a saturation around the value of 0·5. That is, the difference in MAPE reduction is very small when α takes a value beyond 0·5. Based on the above results, the optimum weights to be assigned for the inputs are selected as 0·8 for W1, W2 average section travel time, 0·2 for PV1, PV2, PV3 average section travel time, and the optimum α-value as 0·5. These values were used for predicting the arrival time of buses for the remaining 90 trips in four days, which were kept for validation. The final results of MAPE between the observed and predicted arrival times for 21 bus stops for the remaining 90 trips using the optimum weights and α-values are shown in Figure 3.

It can be seen from Figure 3 that the average MAPE for the four days is 13·25, 13·18, 11·35 and 14·19 and the overall average is observed to be 12·99. From Figure 3 it can also be observed that at bus stop 1, the MAPE of arrival time is very high and after that, it decreases towards downstream bus stops until bus stop 21. The reason for the high MAPE in bus stop 1 is the smaller number of data points available for prediction initially. As the bus moves forward, further data points from more 100 m sections will be available and prediction accuracy increases. Hence, use of a simple historic average is adopted for improving the performance for the first 15 sections of 100 m in length (10% of the total number of sections). The improved results of applying the historic average (simple average of W1, W2, PV1, PV2, PV3 section travel times) for the initial sections are shown in Figure 4, where it can be observed that at bus stop 1, there is a considerable reduction in the MAPE when compared to that of

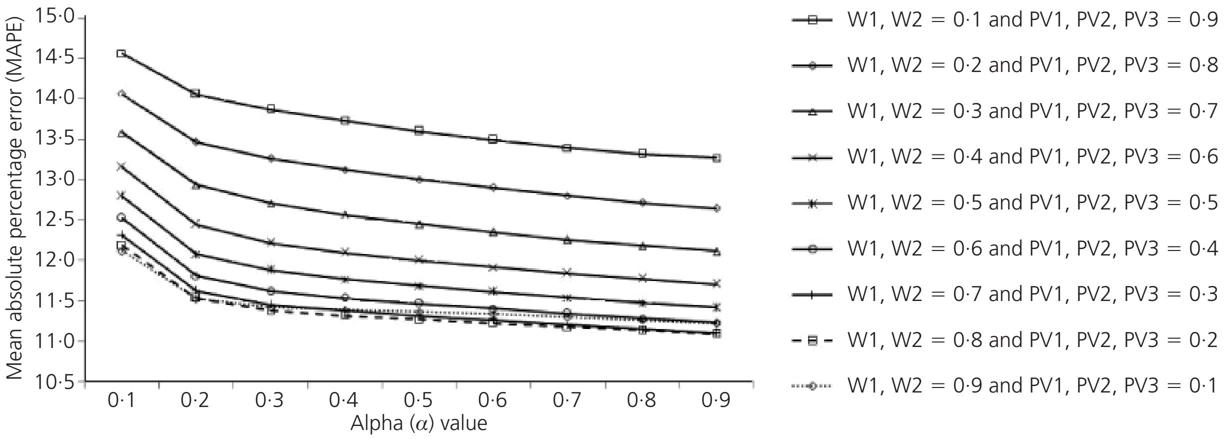

Figure 2. MAPE for varying weights for inputs and varying alpha (α) values

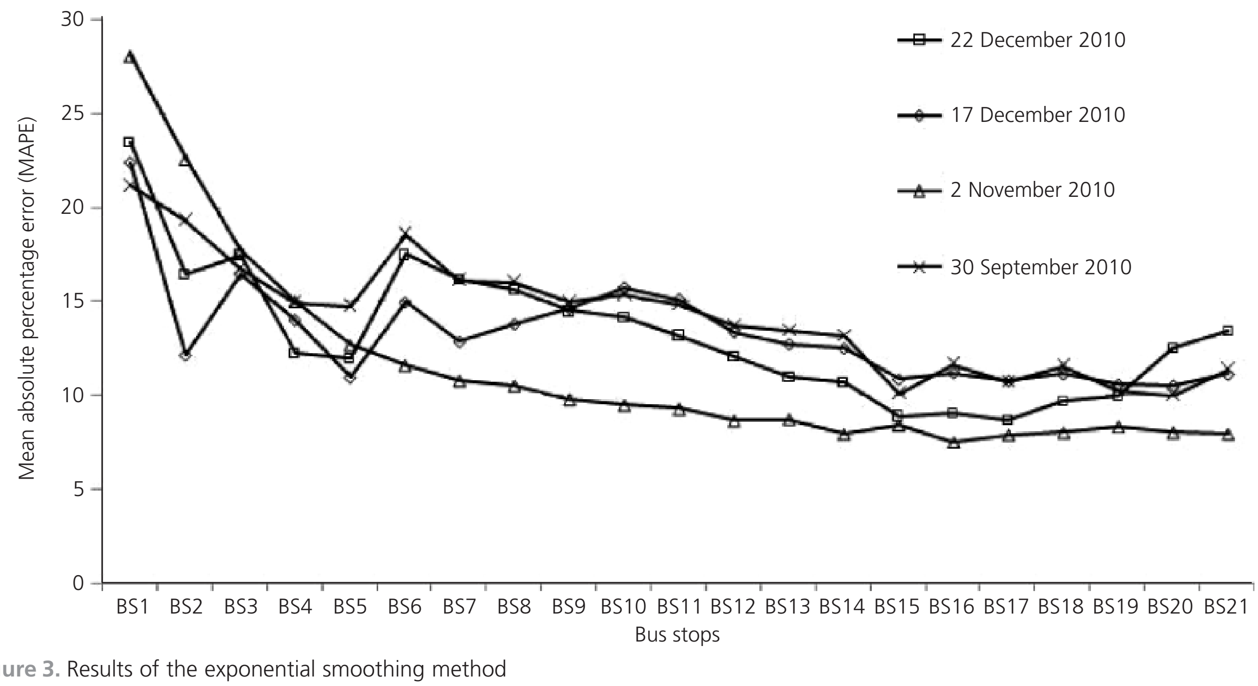

Figure 3. Results of the exponential smoothing method

Figure 3. The average MAPE for the four days from Figure 4 is observed to be 11·24, 11·86, 11·45 and 14·35 and the overall average is 12·22.

According to Warman (2003), passengers have a reasonably high tolerance when the disparity between produced and actual arrival time is under 5 min and if 88% of predicted times were under this. The tolerance dropped off rapidly if the next bus was more than 5 min later than shown on the display. Considering $\pm$ 5 min as an allowable error limit (maximum difference of predicted arrival time with respect to observed/actual arrival time), the number of times the deviation was less than $\pm$1 min, less than $\pm$2 min, less than $\pm$3 min, less than $\pm$4 min, less than $\pm$5 min and more than the tolerable limit $\pm$5 min was found for all the 90 trips and are shown in Figure 5, where it can be seen that on average 77% of the times the deviation is less than the user acceptable range of $\pm$5 min.

In order to check the performance of the exponential smoothing model, it was compared with a base method called the previous trip method, where the travel time of PV3 (the immediate previous vehicle) is assumed to remain the same for

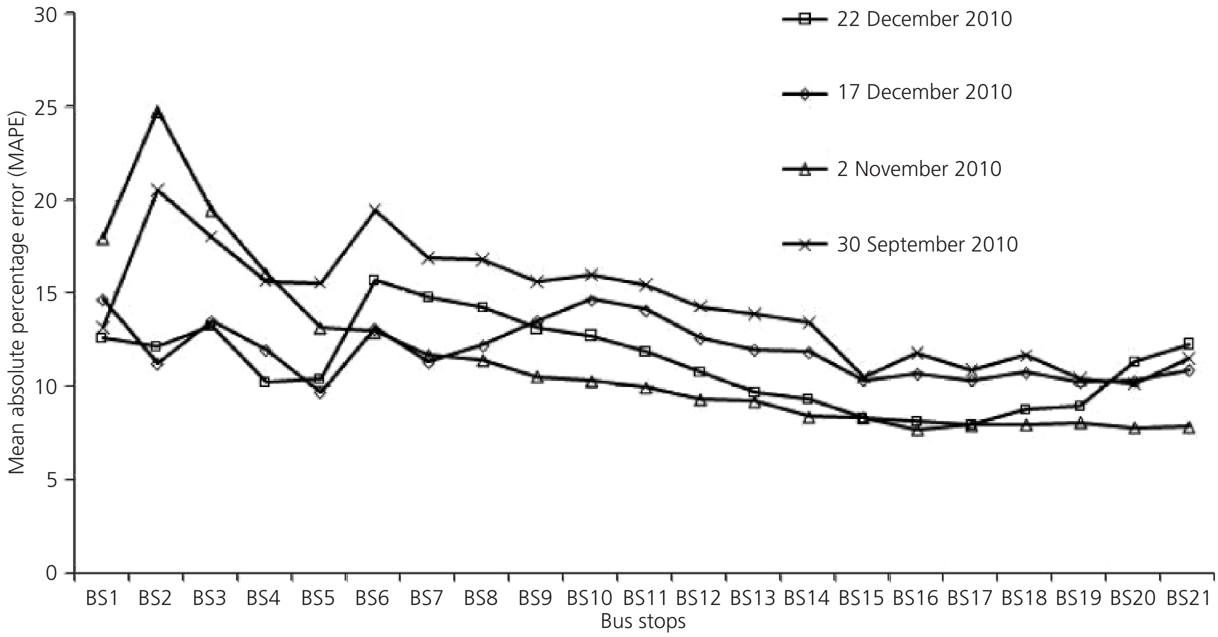

Figure 4. Results of the historic average and exponential smoothing combined model

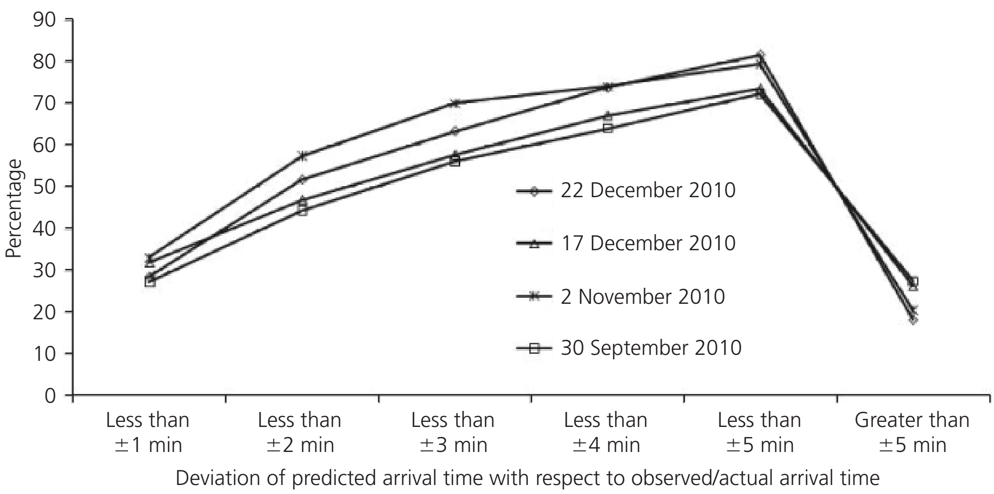

Figure 5. Results of the predicted arrival time deviation

the test vehicle (TV). The results of the comparison of base method with the exponential smoothing model in terms of MAPE between observed and predicted arrival time for the 90 trips of four days is shown in Figure 6. Except for 17 December 2010, on all other days, the exponential smoothing performs better compared to the base method. The results of the percentage number of times the predicted arrival time deviated with respect to observed/actual arrival times for various time categories, for both the base method and exponential smoothing, is shown in Table 1, where it can be seen that, except on 17 December for the less than ± 4 min category and less than ± 5 min category, the exponential smoothing is performing better by exhibiting a higher percentage when compared to the base method. Even for 17 December 2010, for the ± 1 min and ± 2 min categories, the exponential smoothing outperforms the base method, clearly indicating the developed model using exponential smoothing to be better than the base method considered.

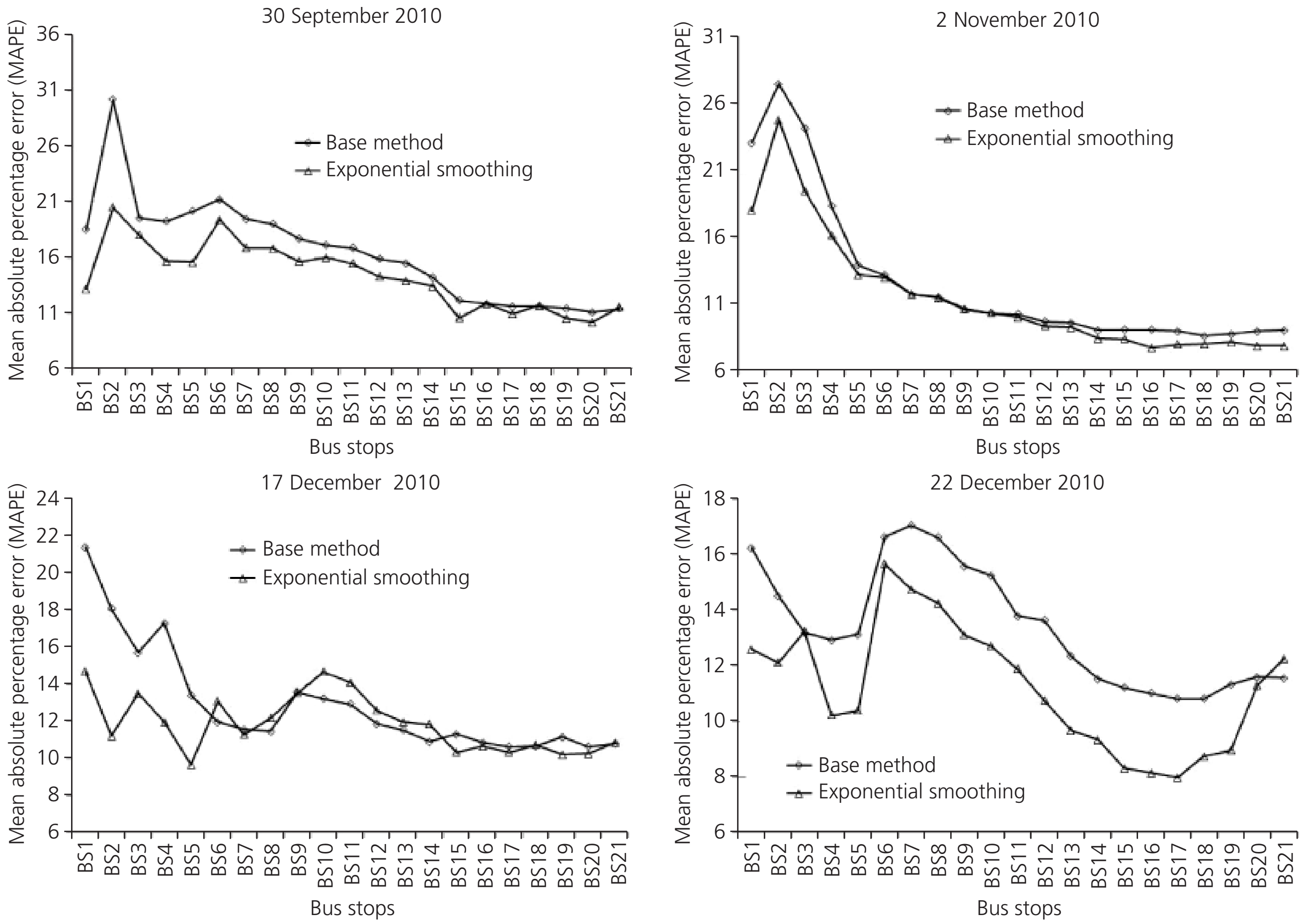

Figure 6. Comparison of exponential smoothing with base method

	30 September 2010		2 November 2010		17 December 2010		22 December 2010	
	Base method	Exponential smoothing	Base method	Exponential smoothing	Base method	Exponential smoothing	Base method	Exponential smoothing
$< \pm 1$ min	17·81	27·33	26·28	32·98	21·43	31·82	21·96	28·57
$< \pm 2$ min	30·02	44·51	47·97	57·5	39·83	46·97	38·62	51·85
$< \pm 3$ min	42·65	56·31	63·84	70·19	56·71	57·79	51·85	63·49
$< \pm 4$ min	57·35	64·18	73·54	74·25	68·18	67·32	68·25	73·81
$< \pm 5$ min	67·29	72·46	80·25	79·72	78·14	73·81	76·72	81·75
$> \pm 5$ min	32·71	27·54	19·75	20·28	21·86	26·19	23·28	18·25

Table 1. Percentage number of times the result deviated for exponential smoothing and base method

5. Conclusion

Advanced public transportation systems (APTS) aim at encouraging public transit usage, thereby reducing congestion on urban roads. Ways to make public transit more attractive are of great importance and one way is to provide real-time status information for public transit buses and expected arrival time of the next service pre-trip and en-route. The reliability of the information provided is very important and is directly dependent on the input data and the prediction model used. The present study tried to find the most significant inputs for predicting the next bus arrival time using valid statistical tests. The study also helped to find out the trip-wise, daily and weekly patterns of transit travel time under heterogeneous traffic conditions. It was found that the previous two weeks' same-day/same-time trips and previous three trips of the same day are most significant for predicting the next bus arrival time. Using these identified inputs, a prediction model was developed using a simple exponential smoothing technique. The optimum relative weights to be assigned for the inputs of previous two weeks' same-day/same-time trips and previous three trips of the same day were found along with the optimum α-value for use with the model. The performance of the smoothing method was corroborated using 90 actual bus trip data spanned over four days in a study route representative of a typical urban environment. The results show that, using significant trips as input, the exponential smoothing method performs better than the previous trip method with an average MAPE of 12. The proposed method was able to predict the bus arrival time with 77% of the times the prediction error being within the user acceptable range of ± 5 min, and 50% of the times within ± 2 min. The results show that, with constraints such as a limited database, technical understanding and transferability, this classical time series method may be considered for real-time APTS implementations.

Acknowledgement

The authors acknowledge the support for this study as part of the project CIE/10-11/168/IITM/LELI by the Ministry of Urban Development, Government of India, through letter no. N-11025/30/2008-UCD.

REFERENCES

Arsham H and Shaw SP (1985) Seasonal and cyclic forecasting for the small firm. *American Journal of Small Business* **9(4)**: 46–57.

Bhandari RR (2005) *Bus Arrival Time Prediction Using Stochastic Time Series and Markov Chains*. PhD thesis, Department of Civil Engineering, New Jersey Institute of Technology, Newark, NJ, USA.

Bin Y, Zhongzhen Y and Baozhen Y (2006) Bus arrival time prediction using support vector machines. *Journal of Intelligent Transportation Systems* **10(4)**: 151–158.

Binoy (2011) *ARIMA vs. Exponential Smoothing*. See http://binoybnair.blogspot.in/2011/09/arima-models-box-jenkins.html (accessed 04/07/2012).

Bo Y, Jing L, Bin Y and Zhongzhen Y (2009) An adaptive bus arrival time prediction model. *Proceedings of the 8th International Conference of Eastern Asia Society for Transportation Studies, Surabaya, Indonesia*. See http://www.easts.info/publications/journal_proceedings/journal2010/100064.pdf (accessed 26/09/2012).

Booth H and Tickle L (2008) *Mortality Modelling and Forecasting: A Review of Methods*. Australian Demographic and Social Research Institute, The Australian National University, Canberra ACT, Australia, Working Paper No. 3.

Chien SI, Ding Y and Wei C (2002) Dynamic bus arrival time prediction with artificial neural networks. *Journal of Transportation Engineering* **128(5)**: 429–438.

Chien SI and Kuchipudi CM (2003) Dynamic travel time prediction with real-time and historic data. *Journal of Transportation Engineering* **129(6)**: 608–616.

Dailey DJ (1999) *An Algorithm for Predicting the Arrival Time of Mass Transit Vehicles using Automatic Vehicle Location Data*. Transportation Research Board, National Research Council, Washington, DC, USA, CD-ROM.

Damrong P and Churueang P (2005) Monthly energy forecasting using decomposition method with application of seasonal ARIMA. *Proceedings of the 7th IEEE International Power Engineering Conference, Singapore*, pp. 1–6.

Gardner ES and Saiz JD (2002) Seasonal adjustment of inventory demand series: a case study. *International Journal of Forecasting* **18(1)**: 117–123.

Jeong RH and Rilett LR (2005) *Bus Arrival Time Prediction Model for Real Time Applications*. Transportation Research Board, National Research Council, Washington, DC, USA, CD-ROM.

Jeong RH and Rilett LR (2004) Bus arrival time prediction using artificial neural network model. *Proceedings of the 2004 IEEE Intelligent Transportation Systems Conference, Washington, DC, USA*, pp. 988–993.

Mondal DK, Kaviraj A and Saha S (2010) Water quality parameters and fish biodiversity indices as measures of ecological degradation: a case study in two floodplain lakes of India. *Journal of Water Resource and Protection* **2(1)**: 85–92.

Nickerson DM and Madsen BC (2005) Non linear regression and ARIMA models for precipitation chemistry in East Central Florida from 1978 to 1997. *Environmental Pollution* **135(3)**: 371–379.

Patnaik J, Chien S and Bladikas A (2004) Estimation of bus arrival times using APC data. *Journal of Public Transportation* **7(1)**: 1–20.

Ramakrishna Y, Ramakrishna P, Laxshmanan V and Sivanandan R (2006) Bus travel time prediction using GPS data. *Proceedings of the Map India, New Delhi, India*.

Schweiger CL (2003) *Real-time Bus Arrival Information Systems – A Synthesis of Transit Practice*. Transportation Research Board, National Research Council, Washington, DC, USA, Report Number TCRP Synthesis 48.

Shurin JB (2010) Environmental stability and lake zooplankton diversity – contrasting effects of chemical and thermal variability. *Ecology Letters* **13(4)**: 453–463.

Suwardo M, Napiah M and Kamaruddin I (2010) ARIMA models for bus travel time prediction. *Journal of the Institution of Engineers Malaysia* **71(2)**: 49–58.

Temraz HK, Salama MMA and Quintana VH (1996) Application of the decomposition technique for forecasting the load of a large electric power network. *IEE Proceedings Generation*

Transmission and Distribution **143(1)**: 13–18.

Vanajakshi L, Subramanian SC and Sivanandan R (2009) Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses. *IET Intelligent Transportation Systems* **3(1)**: 1–9.

Warman P (2003) *Measured Impacts of Real-Time Control and Information Systems for Bus Services*. Transport Direct, Department for Transport, London, UK.

Wexler L (1996) *Decomposing Models of Demographic Impact on the Environment*. International Institute of Applied Systems Analysis, Laxenburg, Austria, Report Number WP-96851996.

WHAT DO YOU THINK?

To discuss this paper, please email up to 500 words to the editor at journals@ice.org.uk. Your contribution will be forwarded to the author(s) for a reply and, if considered appropriate by the editorial panel, will be published as a discussion in a future issue of the journal.

Proceedings journals rely entirely on contributions sent in by civil engineering professionals, academics and students. Papers should be 2000–5000 words long (briefing papers should be 1000–2000 words long), with adequate illustrations and references. You can submit your paper online via www.icevirtuallibrary.com/content/journals, where you will also find detailed author guidelines.