INTEGRATION OF EXPONENTIAL SMOOTHING WITH STATE SPACE FORMULATION FOR BUS TRAVEL TIME AND ARRIVAL TIME PREDICTION

Selvaraj Vasantha Kumar1, Krishna Chaitanya Dogiparthi2, Lelitha Vanajakshi3, Shankar Coimbatore Subramanian4

1School of Mechanical and Building Sciences, Vellore Institute of Technology (VIT University), Vellore, India
2, 3Dept of Civil Engineering, Indian Institute of Technology Madras, India
4Dept of Engineering Design, Indian Institute of Technology Madras, India

Submitted 31 August 2014; resubmitted 20 December 2014; accepted 11 January 2015; first published online 12 October 2015

Abstract. In recent years, the problem of bus travel time prediction is becoming more important for applications such as informing passengers regarding the expected bus arrival time in order to make public transit more attractive to the urban commuters. One of the popular techniques reported for such prediction is the use of time series analysis. Most of the studies on the application of time series techniques for bus arrival time prediction used Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) models, which are presently not suited for real-time implementation. This is mainly due to the necessity and dependence of ARIMA models on a time series modeling software to execute. Moreover, the ARIMA model building process is time consuming, making it difficult to use for real-time implementations. Alternatively, Exponential Smoothing (ES) methods can be used, as they are easy to understand and implement when compared to ARIMA models. The present study is an attempt in this direction, where the basic equation of ES is used, as the state equation with Kalman filtering to recursively update the travel time estimate as the new observation becomes available. The proposed algorithm of state space formulation of ES with Kalman filtering for bus travel time and arrival time prediction was field tested using 105 actual bus trips data along a particular bus route from Chennai, India. The results are promising and a comparison of the proposed algorithm with ES alone without state space formulation and Kalman filtering has also been performed. An information system based on a webpage for real-time display of bus arrival times has been designed and developed using the proposed algorithm.

Keywords: bus travel time prediction; exponential smoothing; Kalman filtering; global positioning systems; time series analysis; real-time bus arrival information system.

Introduction

In recent years, the problem of bus travel time prediction is becoming more important for applications such as informing passengers regarding the expected bus arrival time in real time to make public transit more attractive to urban commuters. The purpose is to shift the personal vehicle users to public transport, so that the number of private vehicles on the road can be reduced, which can ultimately result in less congestion on urban roads. For example, in the city of Boston, US, the commuters mentioned the improved access to real-time arrival information for buses and trains as reasons to pick public transit over other modes (Dickens 2011). Tang and Thakuriah (2012) found that the provision of real-time bus information increased the bus ridership in Chicago. In the city of Visalia, California, US, people are able to receive an alert on their computer or smart phone when their bus is 5, 15 or even 30 minutes away. The alerts are based on real-time information and resulted in an increase in bus ridership by nearly 18% in a month (Peres 2011). In India, the public transport buses in most of the metropolitan cities are gradually being equipped with Global Positioning System (GPS) instruments that could be used to identify the second-by-second location of the buses in real time. To develop an accurate prediction model that can take the GPS data from buses as inputs, various techniques have been reported such as historic and real-time approaches (Yu et al. 2010a, 2010b), machine learning techniques such as Artificial Neural Network (ANN) (Chien et al. 2002; Shalaby, Farhan 2004;
The proposed algorithm of state space formulation of ES with Kalman filtering for bus travel time and arrival time prediction was field-tested using 105 actual bus trips data along a particular bus route from Chennai, India. As the accuracy of ES method largely depends on the value of the smoothing constant selected for prediction, a procedure has been proposed in the present study to find the optimum smoothing constant along with the optimum weights to be used for the inputs. Finally, an information dissemination system using webpage for real-time display of bus arrival times has been designed and developed using the proposed algorithm.

The data collection involved the Automatic Vehicle Location (AVL) data of 105 bus trips spanned across 5 days collected using permanently fixed GPS units in buses of 5C route. For predicting the travel time/arrival time at bus stops for each of these 105 trips (to be called as Test Vehicle (TV) hereafter), the corresponding previous two weeks' same-day same-time trips (called as W1 and W2 hereafter) and previous three trips of the same day (called as PV1, PV2 and PV3 hereafter) were used as the inputs. The selection of these inputs namely, W1, W2 and PV1, PV2, PV3 for predicting the next bus arrival time was based on statistical tests where these
were found to be the most influencing trips for predicting the next bus travel time (Kumar, Vanajakshi 2013). It is important to mention here that the proposed model using ES and KFT can run with data from just the previous trips alone, if previous weeks’ data are not available. The corroboration 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 all the 105 trips. Hence, the data extraction 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.

1. Estimation Scheme

This section starts with the fundamentals of ES. Next, the bus arrival time prediction model based on ES is explained followed by the state space formulation of ES and the procedure to implement it with KFT.

1.1. Fundamentals of ES

Consider a series of data observed up to and including time \((t - 1)\) and need to forecast the next value of the series, \(x_t\). Let the forecast be denoted by \(\hat{x}_t\). When the observation \(x_t\) becomes available, the forecast error is \((x_t - \hat{x}_t)\). The method of ES predicts for the next time period \((t + 1)\), using the forecast from the previous period \(\hat{x}_t\) and adjusts it using the forecast error \((x_t - \hat{x}_t)\). That is:

\[
\hat{x}_{t+1} = \hat{x}_t + \alpha (x_t - \hat{x}_t),
\]

where \(\alpha\) is a smoothing constant that can take values between 0 and 1. From Eq. (1), it can be seen that the new forecast is simply the previous period forecast plus an adjustment for the error that occurred in the last forecast. When \(\alpha\) has a value close to 1, the new forecast will include a substantial adjustment for the error in the previous forecast. Conversely, when \(\alpha\) is close to 0, the new forecast will include very little adjustment. Eq. (1) can be rewritten as:

\[
\hat{x}_{t+1} = \alpha x_t + (1 - \alpha) \hat{x}_t.
\]

Thus, Eq. (2) can be interpreted as a weighted average of the most recent forecast and the most recent observation. Now, Eq. (2) can be expanded by replacing \(\hat{x}_t\) with its components, as follows:

\[
\hat{x}_{t+1} = \alpha x_t + (1 - \alpha) \{\alpha x_{t-1} + (1 - \alpha) \hat{x}_{t-1}\} = \alpha x_t + \alpha (1 - \alpha) x_{t-1} + (1 - \alpha)^2 \hat{x}_{t-1}.
\]

If this substitution process is repeated by replacing \(\hat{x}_{t-1}\) with its components, \(\hat{x}_{t-2}\) with its components and so on, Eq. (3) results in:

\[
\hat{x}_{t+1} = \alpha x_t + \alpha (1 - \alpha) x_{t-1} + \alpha (1 - \alpha)^2 x_{t-2} + \alpha (1 - \alpha)^3 x_{t-3} + \alpha (1 - \alpha)^4 x_{t-4} + ... + \alpha (1 - \alpha)^t x_1.
\]

Thus, \(\hat{x}_{t+1}\) represents a weighted moving average of all past observations with the weights decreasing exponentially; hence the name ‘exponential smoothing’.

1.2. ES for Bus Travel Time Prediction

The basic equation of ES as shown in Eq. (2) was converted from time domain to space domain for the present problem of bus arrival time prediction as shown:

\[
\hat{x}_k = \alpha x_{k-1} + (1 - \alpha) \hat{x}_{k-1}.
\]

The subscript \(t\) in Eq. (2) denotes discretization over time and subscript \(k\) in Eq. (5) denotes discretization over space. To implement this, the study stretch of 15 km was divided into 100 m sections of uniform length and \(k\) in Eq. (5) represents a 100 m section in space. The notation \(\hat{x}_k\) is the predicted travel time of the TV in the \(k\)-th section, \(x_{k-1}\) is the observed travel time in the \((k - 1)\)-th section, and \(\hat{x}_{k-1}\) is the predicted travel time of the TV in the \((k - 1)\)-th section. In order to find an optimum alpha 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 spanned across 5 days, 15 trips of one sample day were used for finding out the optimum weights and alpha value. A total of 9 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 W1, W2 average and 0.9 for PV1, PV2, PV3 average. 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, 9 sub-cases were considered with varying alpha values ranging between 0.1 and 0.9. The ES model as shown in Eq. (5) was executed for all the 15 trips of this sample day using the 81 cases (9x9) with the relative weights for the inputs and varying alpha values. The Mean Absolute Percentage Error (MAPE) is used as a measure of estimation accuracy and is calculated using:

\[
MAPE_{bs} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{b_{pre} - b_{obs}}{b_{obs}} \right| \times 100,
\]

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 time is shown in Fig. 1. As we can see from Fig. 1, the MAPE gradually decreases with increasing weights for W1, W2 average section travel time and a saturation stage is reached when the weight assigned is 0.8. This clearly shows that, the use of 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. Taking this as the
best option, the analysis to choose the best alpha value was carried out. It can be observed that, the MAPE gradually decreased with increasing alpha values, with the reduction in error being very high initially and reaching saturation around the value of 0.5. Based on the above results, the optimum weights to be assigned for the inputs was selected as 0.8 for W1, W2 average, 0.2 for PV1, PV2, PV3 average, and the optimum alpha value as 0.5. Thus, the model for bus travel time prediction based on ES can be written as:

\[ \hat{x}_k = 0.5x_{k-1} + 0.5\hat{x}_{k-1}, \]  

where:

\[ x_{k-1} = 0.8 \left( \frac{W_{1k-1} + W_{2k-1}}{2} \right) + 0.2 \left( \frac{PV_{1k-1} + PV_{2k-1} + PV_{3k-1}}{3} \right). \]  

1.3. State Space Formulation of ES and Integration with KFT

In general, a state space model consists of two equations (Brockwell, Davis 2013): first is called the state equation, which determines the state \( X_{t+1} \) at time \( t+1 \) in terms of the previous state \( X_t \) and a noise term and can be represented as:

\[ X_{t+1} = \{ F_t \} X_t + \{ W_t \}, \quad t = 1, 2, \ldots, \]  

where: \( \{ F_t \} \) is a sequence of \( v \times v \) matrices and \( \{ W_t \} \) is the process disturbance \( \sim N(0, \{ Q_t \}) \). Sometimes, the state equation may also have an external input. The second equation, called the observation equation, expresses the \( w \)-dimensional observation \( Z_t \) as a function of a \( v \)-dimensional state variable \( X_t \) plus noise. Thus:

\[ Z_t = \{ G_t \} X_t + \{ V_t \}, \quad t = 1, 2, \ldots, \]  

where: \( \{ V_t \} \) is the measurement noise \( \sim N(0, \{ R_t \}) \) and \( \{ G_t \} \) is a sequence of \( w \times v \) matrices and \( \{ W_t \} \) and \( \{ V_t \} \) are uncorrelated. The ES method as given in Eq. (5) can be represented in state space form similar to Eq. (9) and Eq. (10) as given:

\[ X_{k+1} = (1 - \alpha)X_k + \alpha U_k + W_k; \]  

\[ Z_k = X_k + V_k, \]  

where: \( U_k \) in Eq. (11) is the external input, which is the observation or measurement at section \( k \) in the present case. \( U_k \) is taken as the average of previous three trips (PV1, PV2, PV3) of the same day travel time in \( k \)-th section. Since the optimum alpha is 0.5, the Eqs (11–12) can be written as:

\[ X_{k+1} = 0.5X_k + 0.5U_k + W_k; \]  

\[ Z_k = X_k + V_k. \]  

The 0.5 in Eq. (13) is the optimum alpha value calculated in Section 1.2. Once the equivalent state space model is established, the recursive equations of KFT can be used to obtain the a priori and a posteriori travel time estimates. Based on sample calculations, the values of \( Q_k \) and \( R_k \) was found to be as 140 and 40 respectively and were used while implementing the KFT algorithm, the steps of which are explained below:

1) The a priori estimate of the travel time was calculated using:

\[ \hat{x}_{(k+1)} = 0.5x_k + 0.5\hat{x}_{(k)} \]  

The symbol hat ‘\(^\wedge\)’ denotes the estimate, the superscript ‘\(^\wedge\)’ denotes the a priori estimate and the superscript ‘\(^*\)’ denotes the a posteriori estimate. Thus, the variable \( \hat{x}_{(k+1)} \) is the a priori estimate of the predicted section travel time of TV in \( (k+1) \)-th section. The variable \( \hat{x}_{(k)} \) is the a posteriori estimate of the predicted section travel time of TV in \( k \)-th section. Since the predicted TV travel time in the first 100 m section is unknown, the observed TV travel time is taken in the place of \( \hat{x}_{(k)} \) for the first 100 m section. The variable \( x_k \) is the observed travel time in the \( k \)-th section, which is the average of previous three trips (PV1, PV2, PV3) of the same day travel time in \( k \)-th section.

2) The a priori error variance (denoted by \( P \)) was calculated using:

\[ P_{(k+1)} = 0.5P_{(k)} + 0.5 + Q_k. \]  

3) The Kalman gain (denoted by \( K \)) was calculated using:

\[ K_{(k+1)} = P_{(k+1)}^{-1} \frac{1}{P_{(k+1)}^{-1} + R_{(k+1)}}. \]
The equations shown above from steps from 1 to 3 are the ‘time update equations’ and the steps 4 and 5 as shown below are called as the ‘measurement update’ equations.

4) The a posteriori travel time estimate of TV was calculated using:

\[ \hat{x}_{k+1} = \hat{x}_{k+1}^* + K_{k+1} \left( z_{k+1} - \hat{x}_{k+1}^* \right), \]  
\[ \text{Eq. (18)} \]

The variable \( \hat{x}_{k+1}^* \) is the a posteriori estimate of the predicted section travel time of TV in \( (k+1) \)-th section, \( K_{k+1} \) is the Kalman gain calculated in step 3, \( \hat{x}_{k+1} \) is the a priori estimate of the predicted section travel time of TV in \( (k+1) \)-th section and calculated in step 1. The variable \( z_{k+1} \) in Eq. (18) is the average of previous two weeks (W1, W2) same day same time trips travel time in \( (k+1) \)-th section, which is used for correcting the a priori estimate calculated in step 1. The results of optimum weights calculated before showed a higher weightage for the previous week same day same time trip travel time than the previous trips of the same day. Hence, the travel time input from previous two weeks was considered for a posteriori estimate calculation (basically the ‘correction’) in step 4 and the travel time of previous three trips of the same day was considered for a priori estimate calculation in step 1.

5) The a posteriori error variance was calculated using:

\[ P_{k+1}^- = (1 - K_{k+1}) P_{k+1}^* \]  
\[ \text{Eq. (19)} \]

Steps 1 to 5 are executed recursively to obtain the travel time estimates of the TV. A flowchart showing the proposed methodology is shown in Fig. 2.

2. Corroboration of the Estimation Scheme

The proposed algorithm based on ES alone (Eqs (7–8)) and smoothing combined with KFT (Eqs (15–19)) for predicting the next bus travel time/arrival time was corroboration using 105 actual bus trips data and the results are presented in this section. The variables considered for corroboration are the predicted arrival time at 21 bus stops for the TV and the predicted section-wise travel time of TV. The MAPE between observed and predicted arrival times averaged over all bus stops for each day is shown in Fig. 3. It can be seen that, the prediction scheme based on ES integrated with KFT performs better than simple ES in 4 out of 5 days with comparatively lesser MAPE. The t-distribution was used to find the Confidence Intervals (CI) for MAPE of both the methods. From t-distribution table, the t value was found to be 2.776 for 95% confidence and degrees of freedom, \( \nu = 5 - 1 = 4 \). Hence, 95% CI for MAPE of ES method was 12.599 ± 1.923, i.e., one can be 95% confident that the MAPE lies between 11 and 15. For smoothing combined with KFT method, the 95% CI was 11.860 ± 2.397, i.e., one can be 95% confident that the MAPE lies between 9 and 14, which is lower than that of the smoothing method alone.

The MAPE between observed and predicted section-wise travel times of the TV for each of the 105 trips is shown in Fig. 4. It can be observed that the state space formulation of ES and integration with KFT resulted in lesser MAPE when compared to using ES alone in 100 out of 105 trips. This indicates the better performance of the proposed model in section-wise bus travel time prediction when compared to the use of smoothing alone without any state space formulation and KFT. The hypothesis testing, namely ‘two sample z-test for the difference between means’ was also carried out to confirm whether smoothing combined with KFT method performs significantly better than simple ES method. The null hypothesis is \( H_0: \mu_1 \geq \mu_2 \) (MAPE of ES combined with KFT method is higher than that of ES method). The alternative hypothesis is \( H_1: \mu_1 < \mu_2 \) (claim: MAPE of ES combined with KFT method is less than that of ES method). Using the MAPE values shown in Fig. 4, for ES combined with KFT method, the mean and standard deviation were found to be 26.993 and 8.408 respectively. For ES method, the mean and standard deviation are 31.886 and 8.846, respectively. The critical values were –1.65 and –1.29 at 0.05 and 0.1 level of significance respectively. The test statistic was found to be equal to –4.108. Since the calculated test statistic falls in the critical region, \( H_0 \) was rejected at 0.05 level of significance.
(95% confidence level). Thus, it can be concluded that ES combined with KFT method performs better than ES in predicting section-wise bus travel times with MAPE significantly less than that of ES method.

The best performing ES combined with KFT method was also evaluated by checking the deviation of the predicted arrival time from the actual arrival time expressed in terms of user understandable units such as minutes or seconds. Here, it is essential to know the acceptable deviation from the user perspective and the achieved accuracy in earlier studies. Lin and Bertini (2004) reported that if the headway between the buses is in the order of ten to fifteen minutes, users expect high accuracy in prediction in the order of one or two minutes. According to Warman (2003), passengers have a reasonably high tolerance when the disparity between predicted and actual arrival time is under 5 minutes if 88% of predicted times were under this. The tolerance dropped off rapidly if the next bus was more than five minutes later than shown on the display. Wall and Dailey (1999) predicted bus arrival time with less than 12% error (i.e. when the predicted bus arrival time is 15 minutes, 70% of the buses will arrive in between 13 and 17 minutes). The TriMet Transit Tracker System (US) reported a minimum of 2 minutes and a maximum of 4.45 minutes as the 95th percentile arrival estimation error while evaluating the system in seven bus routes (Crout 2007). The transit authority in Singapore claimed that, the real-time bus arrival information system available at 215 bus-stops is accurate with the bus arriving within ±3 minutes from the time specified in 8 out of 10 times (LTA 2009). Chien et al. (2003) reported a deviation of +10 minutes to –10 minutes in one of the bus route in New Jersey, US. According to Rajbhandari (2005), for a bus travel time of 1.5 hours, 5 minutes can be considered as an acceptable level of prediction accuracy.

It is to be noted here that, the above studies on evaluation of prediction accuracy was carried out in places where the schedule time is already available and buses stick to the schedule. In such cases one would expect the predicted arrival time information displayed to be fairly accurate from the view point of user tolerance. However, in places where the schedule time for buses at various bus stops along the route is not known and buses do not stick to the schedule, as in the case of present study, the user tolerance will be much higher. In addition, in the present study, the stochasticity in the system as a whole including bus arrival/travel time is very high. The reason for this can be the heterogeneity and less lane disciplined nature of the traffic, high interaction among the vehicles, and lack of exclusive bus lanes forcing the bus to maneuver with the other vehicles. Considering the above limitations and with the trip time of around 50–70 minutes and headway of 15–30 minutes between the buses during most times of the day, an accuracy of ±5 minutes may be considered as an acceptable limit from the user perspective. Considering ±5 minutes as an allowable error limit, the number of times the deviation went less than ±1, ±2, ±3, ±4, and ±5 minutes was found for all the 105 trips. A total of 2205 absolute error values were used (21 bus stops × 105 trips) and the results are shown in Fig. 5. It can be seen that, 76% of the times the deviation was less than the user acceptable range of ±5 minutes. Also, 70% of the times the deviation was within ±4 minutes; 60% of the times the deviation within ±3 minutes; 47% of the times the deviation within ±2 minutes and 30% of the times within ±1 minute.

The deviations of the predicted arrival times for off-peak and peak bus trips was also found and shown in Fig. 6. Based on observations, the time period from 8 am to 11 am and from 5 pm to 8 pm were considered as peak periods and the remaining time periods were considered as off-peak. Thus, based on the starting times of the 105 trips, the trips were classified as either off peak or peak trips and deviations in predicted arrival times were calculated. A total of 1218 absolute error values (21 bus stops × 58 trips) were used for off-peak scenario and a total of 987 absolute error values (21 bus stops × 47 trips) were used for peak hour scenario respectively and the results are shown in Fig. 6. It can be seen that, 81%
of the instances the deviation was less than the user acceptable range of ±5 minutes in case of off-peak period trips and 70% of the instances the deviation was less than ±5 minutes during peak period trips.

Sample plots of the predicted section-wise travel time values using smoothing combined with KFT method against the actual values for the TV for two representative trips of off-peak and peak are shown in Figs 7–8. It can be seen that the predicted travel times during both peak and off-peak traffic conditions follow the trend of observed values closely showing good performance of the developed model under a wide range of traffic conditions.

In order to check the transferability of the proposed model, bus trip data from a completely different route other than 5C route in Chennai, India was collected and analyzed. The bus route selected for this purpose is 23C in Chennai, which connects the Ayanavaram bus terminus in the north western part of the city and the Tiruvanniyur bus terminus in the southern part. The total route length of 23C is 21 km with 20 bus stops. The GPS data of three consecutive bus trips were collected in 23C route and section travel times were extracted. The first two trips (starting time of the input trips are 12:37 pm and 2:57 pm at Ayanavaram bus terminus) were used as inputs in the prediction algorithm to predict the next bus travel time and arrival time, which started at 3:27 pm at Ayanavaram. Since weekly data (W1, W2) was not available in the selected bus route, only two previous bus trips (PV1 – 12:37 pm trip and PV2 – 2:57 pm trip) were considered as inputs in the prediction algorithm, with PV1 for calculation of a priori estimate and PV2 for a posteriori estimate calculation. The results of predicted and actual section-wise travel times of TV (3:27 pm trip) are shown in Fig. 9. It can be seen from Fig. 9 that the predicted travel times follow the trend of observed values closely at most of the sections thus showing better performance of the developed model. The MAPE obtained was 29.22 and 21.99 for travel time and arrival time respectively. The reason for slightly higher MAPE is that, only previous two buses were considered as input here, whereas in case of 5C route, both weekly and previous trip data were considered. The above results of the evaluation of the proposed algorithm in a different bus route in Chennai showed that transferability of the model to other bus routes is not an issue.
3. Web Based Prototype Development

Providing bus arrival details through a website has been rapidly gaining popularity as a medium of information dissemination to the commuters. Web-based travel information dissemination is based on an interactive website where one can select the desired route and bus stop to get the real-time bus arrival and location details. Through the website, a commuter can check the arrival details from the comfort of one’s office or home, and can reach the bus-stop close to the predicted arrival time. Once this information is available to the commuter, it saves a great deal of time for the commuter, who can spend the intermediate time in a more productive way than waiting at the bus-stop. Using the best performing ES combined with KFT method, a web application was developed in the present study to inform commuters the Expected Arrival Time (ETA) at any chosen bus-stop in 5C route. The Google Maps-based web site as shown in Fig. 10 has been developed using languages PHP and MySQL with Ajax support. The website provides the flexibility of choosing the bus-stop at which the user wishes to board the bus.

![Fig. 10. Website showing the bus location and ETA](image)

Concluding Remarks

The bus travel time prediction is an important component in bus arrival prediction applications such as providing accurate bus arrival time information using Variable Message Signs (VMS) or web pages in order to attract more public transit users, which ultimately help to reduce congestion on the urban roads. The existing studies which use time series techniques such as Box-Jenkins ARIMA models suffer certain disadvantages such as the necessity and dependence on a time series modelling software, requirement of sound database for model building, time consuming model building process, more computational time for running the model, difficulty in understanding the model parameters and forecasting principles. These limitations restrict the applicability of Box-Jenkins models for applications, which involve real-time data handling such as the bus travel time and arrival time prediction. As an alternate to ARIMA models, the present study developed a model based on ES and KFT for bus travel time/arrival time prediction. The model was evaluated using 105 actual bus trips data in a typical bus route in Chennai, India. The result showed that the ES combined with KFT performed better than using ES alone for the prediction of bus arrival time at bus stops. Prototypes were also developed using this algorithm for the field implementation of a web based bus arrival prediction dissemination system.

Acknowledgements

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

References


