TRANSPORT
ISSN 1648-4142 / eISSN 1648-3480
2021 Volume 36 Issue 3: 221–234
https://doi.org/10.3846/transport.2021.15220

BUS TRAVEL TIME PREDICTION USING SUPPORT VECTOR MACHINES FOR HIGH VARIANCE CONDITIONS

Anil Kumar BACHU[1], Kranthi Kumar REDDY[2], Lelitha VANAJAKSHI[3*]

[1]*Dept of Civil and Environmental Engineering, Indian Institute of Technology Patna, Bihta, India*
[2, 3]*Dept of Civil Engineering, Indian Institute of Technology Madras, Chennai, India*

Submitted 13 January 2017; resubmitted 24 July 2017, 25 August 2017; accepted 24 November 2017

Abstract. Real-time bus travel time prediction has been an interesting problem since past decade, especially in India. Popular methods for travel time prediction include time series analysis, regression methods, Kalman filter method and Artificial Neural Network (ANN) method. Reported studies using these methods did not consider the high variance situations arising from the varying traffic and weather conditions, which is very common under heterogeneous and lane-less traffic conditions such as the one in India. The aim of the present study is to analyse the variance in bus travel time and predict the travel time accurately under such conditions. Literature shows that Support Vector Machines (SVM) technique is capable of performing well under such conditions and hence is used in this study. In the present study, nu-Support Vector Regression (SVR) using linear kernel function was selected. Two models were developed, namely spatial SVM and temporal SVM, to predict bus travel time. It was observed that in high mean and variance sections, temporal models are performing better than spatial. An algorithm to dynamically choose between the spatial and temporal SVM models, based on the current travel time, was also developed. The unique features of the present study are the traffic system under consideration having high variability and the variables used as input for prediction being obtained from Global Positioning System (GPS) units alone. The adopted scheme was implemented using data collected from GPS fitted public transport buses in Chennai (India). The performance of the proposed method was compared with available methods that were reported under similar traffic conditions and the results showed a clear improvement.

Keywords: support vector machines, bus travel time prediction, approximate entropy, high variability, heterogeneous traffic.

Notations

ANN – artificial neural network;
ApEn – approximate entropy;
APTS – advanced public transportation system;
ATIS – advanced traveller information system;
COV – coefficient of variation;
GPS – global positioning system;
HMHV – high mean high variance;
ID – identification;
IIT – Indian Institute of Technology;
IT – information technology;
KFT – Kalman filtering technique;
LMLV – low mean low variance;
MAE – mean absolute error;
MAPE – mean absolute percentage error;
MLR – multiple linear regression;
MSE – mean square error;
MTC – Metropolitan Transport Corporation (Chennai) Ltd (*https://mtcbus.tn.gov.in*);
NN – neural network;
SIPCOT – State Industries Promotion Corporation of Tamil Nadu (*https://sipcot.tn.gov.in*);
SRM – structural risk minimization;
SVM – support vector machines;
SVR – support vector regression.

Introduction

One of the key elements in ATIS and APTS is to predict vehicle travel time or arrival time with reasonable accuracy. This would have been an easy task if the deviations in travel times were minimal or having less uncertainty. However, these deviations are mainly due to traffic signals, congestion, and weather condition, where the uncertainty is high. The variability and uncertainty in travel time is much higher in a heterogeneous and lane-less traffic condition such as the one existing in India. Manually

*Corresponding author. E-mail: *lelitha@iitm.ac.in*

operated signals and presence of pedestrians and animal drawn carts in the traffic stream adds to these complexities. Developing a model that can take into account all these factors is a challenging task. Existing literature mainly focused on three types of models for travel time prediction: time series (Rajbhandari 2005; Suwardo *et al*., 2010; Kumar, Vanajakshi 2012), ANN (Jeong, Rilett 2004; Ramakrishna *et al.* 2006; Kumar *et al.* 2014b; Fan, Gurmu 2015; Chen *et al.* 2007; Vanajakshi, Rilett 2004; Mazloumi *et al.* 2011), and KFTs (Liu *et al.* 2014; Nanthawichit *et al.* 2003; Shalaby, Farhan 2004; Vanajakshi *et al.* 2009; Padmanaban *et al.* 2010; Kumar, Vanajakshi 2014; Chu *et al.* 2005; Kumar *et al.* 2014a). There were only a few studies that paid special attention to high variability issue (Mazloumi *et al.* 2011) using ANN. None of the studies paid special attention to address the high variability issue under heterogeneous and lane-less traffic conditions that are leading to higher prediction errors on certain sections and trips. This may be because the model equations used in those studies were developed based on simple equations for characterizing the evolution of travel time. SVM has been reported as a forecasting tool that can perform well with uncertainty in several areas in recent years (Yu *et al.* 2006, 2010a; Wu *et al.* 2003; Vanajakshi, Rilett 2004, 2007). SVM is a specific type of learning algorithm characterized by the use of kernel functions. SVM theory mainly depends on SRM principle to estimate a function by minimizing an upper bound of the generalization error. SVM is shown to be very resistant to over fitting problem (Vapnik 1999). Another important property of the SVM is that training SVM is equivalent to solving a linearly constrained quadratic programming problem so that the solution of SVM is always unique and globally optimal, unlike other network training that requires nonlinear optimization, which may lead to local minima. Table 1 presents a summary of the studies that used SVM to predict bus travel time along with the corresponding traffic conditions and variables considered.

In the present study, a reliable system for real-time bus travel time prediction paying special attention to the high variability condition was developed using nu-SVR. The optimum amount of data required to predict the next trip was identified by ApEn technique and performance of the proposed method was compared with a model based approach (Vanajakshi *et al.* 2009) that was reported to be performing well under Indian traffic conditions. The validation was done for a selected bus route in Chennai (India), which are equipped with GPS.

1. Literature review

Various techniques have been reported in literature for the prediction of travel time. These include approaches such as historical and real-time averaging, statistical analyses (Yu *et al.* 2017; Bian *et al.* 2015; Xu, Ying 2017), dynamical systems approach and machine-learning techniques. Each of these techniques has its own advantages and disadvantages in terms of data requirement, variables involved and complexity of analysis. Historical and real-time averaging methods are sufficient under expected traffic conditions without much variation. However, under unexpected traffic conditions where the variability is high, their accuracy will be greatly reduced (Jeong, Rilett 2004; Shalaby, Farhan 2004; Vanajakshi, Rilett 2007). Regression methods will predict the dependent variable (travel time) using a set of independent variables that can affect travel time (Abdelfattah, Khan 1998; Patnaik *et al.* 2004; Kwon *et al.* 2000; Ramakrishna *et al.* 2006; Yu *et al.* 2010b). However, it is difficult to identify and collect information on the exhaustive set of affecting variables (Cheng *et al.* 2010). Dynamical systems approaches develop models that can capture the dynamics of the system by establishing mathematical relationships between appropriate variables (Wall, Dailey 1999; Dailey *et al.* 2001; Cathey, Dailey 2003; Shalaby, Farhan 2004; Vanajakshi *et al.* 2009; Padmanaban *et al.* 2010; Kumar, Vanajakshi 2012, 2014; Hans *et al.* 2015; Zhou *et al.* 2017; Kumar *et al.* 2017). However, it may not be always possible to develop explicit equations that can capture the system dynamics very efficiently. Machine learning techniques such as ANN (Chen *et al.* 2004; Ramakrishna *et al.* 2006; Jeong, Rilett 2004; Mazloumi *et al.* 2011) and SVM (Yu *et al.* 2006, 2016; Wu *et al.* 2003; Yang *et al.* 2016) are commonly used to predict travel time because of their ability to solve complex non-linear relation-

Table 1. Summary of literature that used SVM for bus travel time prediction

Source	Traffic conditions	Variables considered
Proposed method	heterogeneous	travel time data from previous five vehicles for temporal SVM model and from five previous subsections for spatial SVM models
Yu *et al.* (2006)	homogeneous	time-of-day and weather conditions
Yu *et al.* (2010b)	homogeneous	speed of the bus
Yu *et al.* (2010a)	homogeneous	time-of-day, weather conditions, route segment, travel times on the current segment, latest travel times on the predicted segment
Yu *et al.* (2011)	homogeneous	bus time interval among the route set, bus time interval of the same route, weighted average of bus running time among the route set, bus running time in the same route
Yu *et al.* (2016)	homogeneous	time-of-day, weather conditions and bus speed
Yang *et al.* (2016)	homogeneous	time period, length of road, weather conditions and bus speed

ships. Out of these, SVM has proved to be one of the most effective tools for pattern recognition across different areas and hence is used in this study. A review of reported studies that used SVM for the prediction of traffic parameters are discussed below.

Wu *et al.* (2003) and Vanajakshi, Rilett (2007) used SVM for travel time prediction and showed SVM giving better results compared to historic and real-time methods. Vanajakshi and Rilett (2004) used ANN and SVM to predict the traffic stream speed. The study reported SVM as a viable alternative to ANN for short term prediction of traffic parameters, especially when less data is available for training or when the training data has more variations. Yu *et al.* (2006) predicted the bus arrival time based on the travel time of current segment and the latest travel time of the next segment using SVM. Separate models were built according to the time-of-day and weather conditions. The developed models were tested using off-line data of a transit route and exhibited advantages over an ANN-based model. Yu *et al.* (2010b) developed SVM based models to predict bus arrival time based on speed data. Yu *et al.* (2010a) developed a hybrid model based on SVM and KFT and predicted the baseline travel times based on time-of-day, weather conditions, route segment, travel times on the current segment, and latest travel times on the predicted segment. Yu *et al.* (2011) proposed models to predict bus arrival times based on running times of multiple routes. They developed SVM, ANN, *k*-nearest neighbours algorithm and linear regression models to predict bus arrival time. Yu *et al.* (2016) proposed a model to predict bus travel time using SVM. The arrival time of bus was predicted by taking the time, weather and speed data as input. Yang *et al.* (2016) developed a model to predict bus arrival times using SVM with genetic algorithm. In the study, the characteristics of the time period, the length of road, weather, and bus speed were used as input vectors and genetic algorithm was used to identify the best parameters. From the above discussion, it can be seen that researchers found SVM technique to be producing better results when compared to ANN and other standard techniques for prediction problems. However, no significant studies have been reported using SVM to predict bus travel time under Indian traffic conditions, which differ largely from the conditions in western countries and is the focus of the present study. Studies reported to predict bus travel time under heterogeneous conditions, are discussed below.

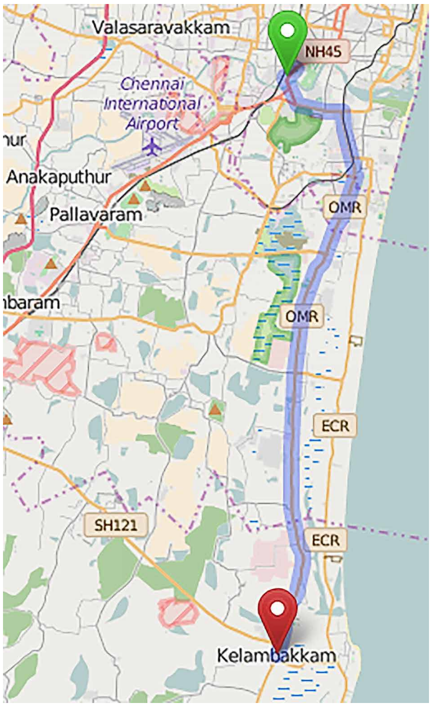
Ramakrishna *et al.* (2006) used NNs to predict the travel time under heterogeneous traffic conditions. They used travel time data from 25 trips of buses to develop ANN and MLR models. Results showed ANN performing better compared to MLR. Vanajakshi *et al.* (2009) used KFT to predict travel time under heterogeneous traffic conditions. They used preceding two bus trips data collected using GPS to predict next bus travel time. Padmanaban *et al.* (2010) extended the above study by incorporating the delays in the model. Kumar and Vanajakshi (2014) proposed a statistical methodology to find out patterns in the data and used them as input to predict the next bus travel time using KFT. Kumar *et al.* (2014a, 2014b) used GPS data to predict bus travel time using ANN and the obtained results were compared with KFT. Results showed that ANN gave better results when there is a large data set for network training. However, many of these studies reported higher errors on sections with high variability such as the ones with signals and bus stops (Fatima, Kumar 2014). This problem of high variability is of bigger concern under the heterogeneous and lane-less traffic. SVM has been reported to perform better when the data is having high variability (Vanajakshi, Rilett 2004) and hence can be a better tool for travel time prediction under such traffic conditions. A few studies were reported from the homogeneous traffic conditions addressing the issue of predicting travel time variability as discussed below.

Fu, Rilett (1998) and Pattanamekar *et al.* (2003) presented a set of analytic functions for both mean and variance of travel times. Liu *et al.* (2005) developed a macroscopic model for urban link travel time prediction based on measurements collected by single loop detectors. The method divided travel times into two components as link cruising times and intersection delays. They then presented a set of analytic equations to derive the mean and the variance of link travel time. Li (2006) developed fuzzy NNs to predict mean travel time and used the S-shaped relationship to predict vehicle-to-vehicle travel time variability. Mazloumi *et al.* (2011) developed two separate ANN models to predict average bus travel time and its variability using traffic flow data, weather conditions and schedule adherence as input variables. The current study is one of the first attempts, where SVM technique is being used for bus travel time prediction for a traffic system with very high variance such as the heterogeneous and lane-less Indian traffic. In addition, the present study is different from the past literature in terms of the variables used in the prediction method. The present study used only the travel time collected using GPS units fitted in buses as input.

2. Data collection and preliminary analysis

Data were collected using GPS units that were fitted in MTC buses in the city of Chennai. The bus route selected for the present study is the Metropolitan Transport Corporation route number 19B, from Saidapet (within city area) to Kelambakkam (in the suburban area). It has a route length of 30 km having 21 major bus stops. The average time headway between two consecutive buses in this route was around 15…30 min. The selected route is shown in Figure 1. This road stretch passes through several types of urban roads with various land use characteristics such as commercial (Saidapet, SIPCOT, etc.), institutional areas (Hindustan Engineering College, Women's Polytechnic College, College of Engineering Guindy, etc.) and IT hubs (Navallur IT Park, Tidel Park, etc.). Sample photographs of the route, 19B at some of the locations are shown in Figure 1b. Under such traffic conditions, a mix of motorized vehicles such as passenger cars, buses, trucks, three wheelers, and two-wheelers would be moving along with non-motorized vehicles such as bicycles and animal drawn carts. All these vehicles share the same road space without any segregation for the various vehicle types as shown in Figure 1b. This will lead to high level of uncertainties and variability in travel time, which make the present study different from the studies that were reported in literature.

a)



b)

Figure 1. 19B bus route: a – route scheme (source: *https://www.openstreetmap.org*); b – snapshots of traffic conditions along bus route at some locations

The collected GPS data includes the ID of the GPS unit, time stamp, and latitude and longitude of the location at which entry was made. From the collected data, the distance between two consecutive entries was calculated using the Haversine formulae (Chamberlain 1996). The processed data consists of travel times and corresponding distance between consecutive locations of all the buses. One month's data were considered for the analysis.

To start with, the entire route length was divided into smaller subsections of 500 m length and the time taken to cover each subsection was calculated using the linear interpolation technique. These data were analysed to find the variations in travel time. Figure 2a shows the variation in travel time across various sections in the considered study route and it can be observed that a few sections are experiencing high variability in travel time. To understand the reason for these high values, a closer look of the route was carried out and was observed that each of the peaks corresponded to a bus stop or intersection. Table 2 presents those details (column titled "Characteristics") along with their corresponding descriptive statistics such as average travel time, average speed, standard deviation and COV.

Next, variations in travel time over time of the day were analysed. For this, travel times experienced by all trips in individual sections were studied. Figure 2b and Figure 2c show, sample plots, one from subsection 7, a low variance subsection from the suburban area and other from subsection 47, a high variance subsection from the urban area. From Figure 2b, it can be observed that the travel times during 8:00…10:00 AM and 4:00…6:00 PM are relatively high, indicating morning and evening peak in traffic. In addition, it can be observed that the peak hours have more variance than those in off-peak hours. On the other hand, Figure 2c shows very little variation in travel time in this subsection over different time periods of the day. The main reason for this can be that the subsection is in the suburban area, where the traffic is less and the effect of peak and off-peak is not very prominent compared to the ones in urban areas (subsection 46) as shown in Figure 2b.

A similar analysis was carried out to study the effect of rain on bus travel time. Data were collected on 7 rainy days and the travel times were compared with the travel time of sunny days. Figure 2d shows the effect of rain on bus travel time for sample subsections. It can be seen that both mean and variance of travel time are higher on rainy days compared to sunny days.

Thus, it can be seen that sections in urban areas during peak hours and the sections with intersections or other obstructions experience more travel time and variability. In addition, weather changes such as rain causes variation in travel time.

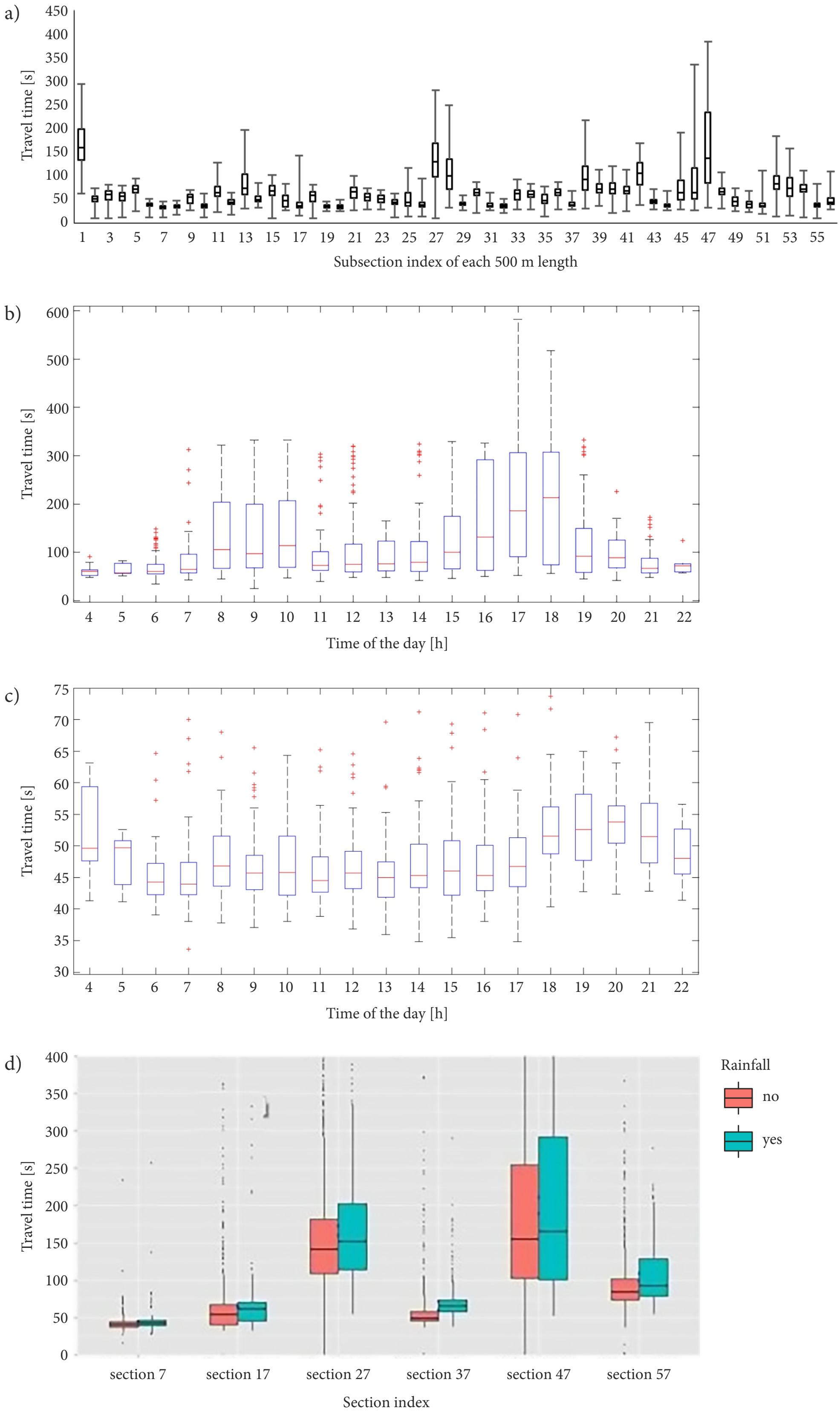


Figure 2. Variations: a – variation in travel time across various sections along the route; b – travel time variation in subsection 46 (urban); c – travel time variation in subsection 7 (suburban); d – effect of rainfall on bus travel time for sample subsections

Table 2. Descriptive statistics of high mean and variance subsections

Subsection ID	Section length [km]	Locations	Characteristics	Average travel time [s]	Average speed [km/h]	Standard deviation [s]	COV [%]
11	0.5	Toll Booth	toll booth	74	24.14	43	57.26
13	0.5	SIPCOT	bus stop	92	19.42	52	55.84
17	0.5	Navullur Church	bus stop	92	19.50	62	67.32
26…27	1.0	Sholinganallur	bus stop and 4-legged signalized intersection	192	18.69	92	47.81
28	0.5	Accenture	Bus stop	118	15.22	70	59.69
37…38	0.5	Ellaiamman Nagar	3-legged skewed intersection	152	23.66	79	52.17
40…42	1.5	Thoraipakkam	"T" intersection	260	20.72	173	66.36
45…46	1.0	SRP tools	"T" intersection	196	18.34	155	78.90
47	0.5	Tidel park	bus stop and 2 major intersections within 100 m distance	171	10.50	109	63.86
51…52	1.0	Kotturpuram (IIT)	"T" shaped signalized intersection	155	23.21	137	88.37

3. SVM

From literature review, it was observed that a few studies predicted the variability in travel time separately using ANN and the total travel time using SVM under homogeneous traffic conditions, and none of the studies under heterogeneous traffic conditions. Earlier studies (Vanajakshi, Rilett 2004) stated SVM as a viable alternative to ANN for prediction of traffic parameters, when less amount of data is available for training or when the training data has more variations. Considering the high variation in travel time, particularly with the heterogeneous and lane-less traffic conditions, SVM seems to be a suitable candidate for prediction. The present study tries to predict bus travel time especially under these conditions using SVM and the details are presented below.

SVM are learning systems that use a hypothetical space of linear functions in a high dimensional feature space, trained with a learning algorithm. This learning strategy was introduced by Vapnik (1999). The main idea behind SVM is that for a given training sample, the SVM construct a hyper plane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. SVM are based on the SRM inductive principle, which seeks to minimize an upper bound of the generalization error consisting of the sum of the training error and confidence level. In SVR, the basic idea is to map the data into high dimensional feature space via non-linear mapping and do linear regression in this space. This linear regression in high dimension space corresponds to non-linear regression in the low dimension input space.

Consider a set of training data points $(x_1, y_1), (x_2, y_2)$, …, (x_n, y_n), where x_n is an n-dimensional input vector such as previous travel times of current segment, and y_n is the desired value. Let $\hat{y}_n$ be the predicted value such as travel time of next segment, and n is the number of training samples. Let the output data vector be in the form:

$$y = f(x). \tag{1}$$

SVM approximates the function in Equation (1) using the following form:

$$\hat{y}(\bar{x}, \bar{\omega}) = \sum_{i=1}^{n} \omega_i \cdot \phi_i(x) + \omega_0 = \bar{\omega}^t \cdot \phi(x) + \omega_0, \tag{2}$$

where: $\phi(x)$ represents the high dimensional feature spaces, which is nonlinearly mapped from the input space x; the coefficients $\omega_0, \bar{\omega}$, etc. are estimated by solving a constrained optimization problem, which is done using Lagrangian multiplier method. The regression problem can be solved by using nu-SVR, which was proposed by Schölkopf *et al.* (2006). The basic optimization equations for nu-SVR is:

cost function:

$$\frac{1}{2} \cdot \|w\|^2 + c \cdot \nu \cdot \varepsilon + \frac{c}{N} \cdot \sum_{n=1}^{N} (\xi_n - \xi'_n) \tag{3}$$

with constraints as:

$$\begin{aligned} & y_n - \hat{y}_n = \varepsilon + \xi_n; \\ & \hat{y}_n - y_n = \varepsilon + \xi'_n; \\ & \varepsilon_n \geq 0; \\ & \xi'_n \geq 0; \\ & \varepsilon > 0, \end{aligned} \tag{4}$$

where: w – weights vector; c – regularization constant; ν – lower bound on the fraction of support vectors (a scalar bounded in between 0 and 1); ε – width of the tube; ξ_n, ξ'_n – slack variables.

In dual form, the above can be represented as:

$$L_d(\bar{\alpha}, \bar{\alpha}') = \sum_{n=1}^{N} \varepsilon_n \cdot (\alpha_n - \alpha'_n) - \frac{1}{2} \cdot \sum_{m=1}^{N} \sum_{n=1}^{N} (\alpha_m - \alpha'_m) \cdot (\alpha_n - \alpha'_n) \cdot k \cdot (x_m, x_n) \tag{5}$$

with constraints as:

$$\sum_{n=1}^{N} \varepsilon_n \cdot \left(\alpha_n - \alpha_n'\right) = 0;$$
$$0 < \alpha_n \cdot \alpha_n' \leq \frac{C}{N};$$
$$\sum_{n=1}^{N} \left(\alpha_n - \alpha_n'\right) \leq c \cdot \nu, \tag{6}$$

where: L_d – Lagrangian multiplier in dual form; $\bar{\alpha}$, $\bar{\alpha}'$ – Lagrangian multipliers; "–" – represents vector; $m = 1, ..., N$, $n = 1, ..., N$ – denote N-dimensional vectors. ε can be imagined as a tube, equivalent to the approximation accuracy placed on the training data points. If the predicted values are within this accuracy limit, the loss associated with that point is assumed as zero, and if the predicted point is outside this accuracy limit, the loss is taken as the magnitude of the difference between the predicted value and radius ε of the tube. A large ε can depreciate the approximation accuracy placed on training points. C is called the regularization constant. Increasing the value of C will result in the relative importance of the empirical risk with respect to the regularization term. Both C and ε are user prescribed parameters. In this study, to predict the travel time for the next instances, LIBSVM tool box in MATLAB (*https://www.mathworks.com/products/matlab.html*) was used (Chang, Lin 2011), and the kernel function used was a linear kernel of the form:

$$K\left(x_i, x_j\right) = \gamma\left(x_i \cdot x_j\right) + coef, \tag{7}$$

where: K – linear function; γ – width parameter; *coef* – coefficient.

4. Implementation and results

4.1. Model development using SVM

In the present study, two different prediction approaches were attempted using SVM, namely temporal and spatial. The temporal approach used the travel times of previous many trips of same subsection to predict next trip in that subsection, whereas the spatial approach used travel time of the same trip on previous many sections to predict next subsection travel time. Unlike earlier studies, the present study used only the travel time data collected using GPS units fitted in buses for the prediction. The details of input selection and SVM implementation are discussed separately in the sections below.

4.1.1. Prediction model considering spatial variation (spatial SVM)

Here, bus travel time was predicted by considering the spatial variation in travel time, where the travel time of the same trip on previous many subsections is used to predict the travel time of the next subsection. In order to identify the optimum inputs (number of previous subsections), ApEn technique was used. ApEn is a technique used to quantify the amount of regularity and the unpredictability of fluctuations in data over time (Pincus 1991). This

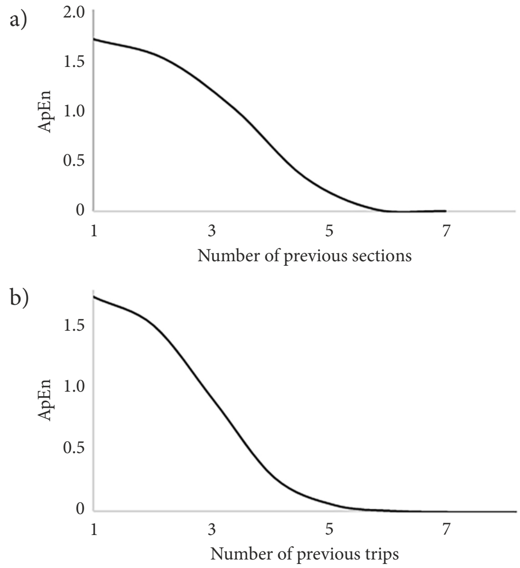


Figure 3. ApEn: a – vs number of previous sections; b – vs number of previous trips

algorithm was implemented for the current study to find the uncertainty associated with predicted travel time and was implemented using MATLAB. ApEn obtained for the training data with respect to number of previous sections is shown in Figure 3a. It can be observed in the figure that the uncertainty in the prediction is negligible after using travel time from previous five sections as input. Hence, previous five subsections travel time was selected as the input to predict the bus travel time in the next subsection. Thus, input vector to SVR was a five-dimensional matrix with previous five subsections travel times, and output vector consisted of corresponding next subsection travel time. Models were developed for each subsection separately using these identified inputs.

4.1.2. Prediction model considering temporal variation (temporal SVM)

In the case of temporal SVM, bus travel time was predicted by considering the temporal variation in travel time, meaning the travel time of the next trip was predicted using the travel time of previous trips in the same subsection. In this case, also, ApEn was used to find the optimum amount of data required to predict the next trip. Figure 3b shows the ApEn obtained for the training data with respect to number of previous trips. Here, it can be observed that the uncertainty in the prediction is negligible if previous six or more trips are used as input. Hence, previous six trips was selected as the input size to predict the next bus travel time in this case. Thus, input vector to SVR was a six-dimensional matrix with previous six trips travel time, and output vector consisted of corresponding next trip travel time and was developed for each subsection separately.

SVM requires a good amount of training dataset, with inputs and the corresponding outputs. One month's data were used in this study, out of which 18 days data were used for training, 7 days data for cross-validation and the remaining data to test the performance. The descriptive statistics of these data sets are presented in Table 3 for selected subsections.

During cross-validation, the model is simulated with the separate dataset kept for cross-validation to check the efficiency. During testing, the predicted value obtained from simulation were compared with the actual value. The common measures used for representing forecast accuracy are scale-dependent measures (based on the absolute error or squared errors, e.g. MAE or MSE), and measures based on percentage errors (scale independent, e.g. MAPE). Out of these MAPE is reported as one of the best (Makridakis 1993), and hence was reported in the current study along with correlation coefficient *r*, which are calculated as:

$$MAPE = \frac{\sum_{i=1}^{N} \frac{\left| x_{i,a} - x_{i,p} \right|}{x_{i,a}}}{N} \cdot 100\%\,; \tag{8}$$

$$r = \frac{\sum \left(x_{i,p} - \bar{x}_p \right) \cdot \left(x_{i,a} - \bar{x}_a \right)}{\sqrt{\sum \left(x_{i,p} - \bar{x}_p \right)^2} \cdot \sqrt{\sum \left(x_{i,a} - \bar{x}_a \right)^2}}\,, \tag{9}$$

where: r – correlation coefficient $x_{i,a}$ – actual travel time observed from the field; $x_{i,p}$ – predicted travel time value; $\bar{x}_a$ – average of actual travel time; $\bar{x}_p$ – average of predicted travel time.

Programs were written in MATLAB to generate input vector and output vector for training, validation and testing nu-SVR with linear kernel function was used. The four important unknown parameters of the kernel function were ν (lower bound on the fraction of support vectors), γ (width parameter), *C* (cost/penalty parameter), and *coef* (coefficient), and were obtained by cross-validation.

4.2. Performance evaluation

The results obtained from the implementation of the prediction methods presented in previous section, which will be referred as spatial SVM and temporal SVM, will be discussed in this section. Since the proposed algorithm uses significant historic data as inputs, a comparison was made with the earlier studies that reported the use of model based approaches (named as spatial KFT and temporal KFT). In spatial KFT approach (Vanajakshi *et al.* 2009), the travel time of a bus in an upcoming subsection was predicted using the travel time in the previous subsection. In temporal KFT approach (Kumar *et al.* 2014a, 2014b) the travel time of a bus in a subsection was predicted using previous many buses travel times in the same subsection.

Performance evaluation of the proposed methods was carried out by comparing the predicted values with the actual values over a period of one week for various trips and subsections. Evaluations were carried out separately for peak and off-peak conditions, low and high variance sections (suburban and urban) and rainy and sunny days, based on the findings from the preliminary data analysis and are presented below.

4.2.1. Performance comparison of peak and off-peak conditions

In this scenario, performance comparison was carried out for each of the trips in a day averaged across all sections. Figure 4a shows the performance of the proposed methods (temporal SVM and spatial SVM) for a representative day separated into peak and off-peak. To compare the performance, results using temporal KFT and spatial KFT are also shown in the plot. From Figure 4a, it can be observed that overall, the SVM models having lesser error than the corresponding KFT. On closer look, it can be seen that the temporal SVM performs the best during peak periods (8:00…10:00 AM and 4:00…7:00 PM), and both temporal and spatial SVM performing comparable during off-peak period. Overall, it can be stressed that during peak periods, when the variability is high, temporal SVM outperforms all the other methods.

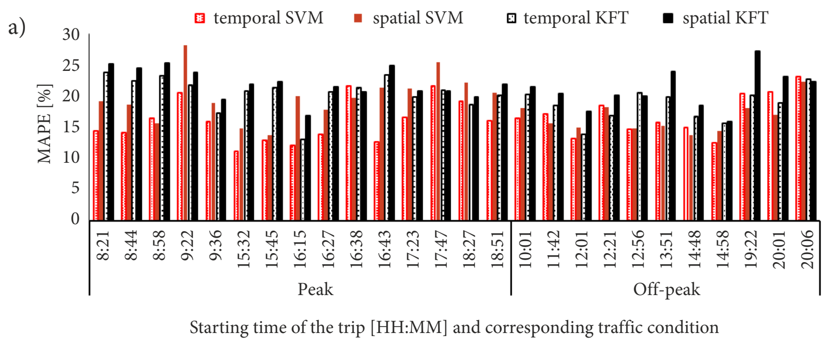
Performance comparison across multiple days was also carried out and Figure 4b shows the results in terms of MAPE for the selected one-week test period. It can be observed here also that temporal SVM outperforms all the other approaches on majority of the days.

4.2.2. Performance comparison over subsections

Along with the comparison of performance over trips, a comparison over subsections was also made. Figure 5 shows a comparison between predicted travel times and actual travel times that were calculated for the maximum

Table 3. Descriptive statistics of the data used for training, validating and testing the SVM model

Subsection ID	Average travel time [s] (average speed [km/h])			Standard deviation [s]			COV [%]		
	training	validation	testing	training	validation	testing	training	validation	testing
11	78 (23.05)	79 (22.93)	74 (24.12)	48	52	52	61.54	65.65	69.35
13	92 (19.48)	93 (19.37)	95 (18.99)	59	55	58	64.29	59.35	61.39
28	116 (15.47)	110 (16.36)	118 (15.23)	88	71	70	75.77	65.03	59.39
47	194 (9.25)	191 (9.43)	204 (11.12)	152	135	148	78.09	70.75	72.65



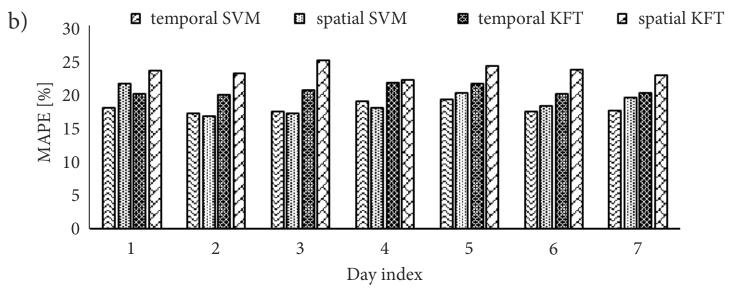


Figure 4. Performance comparison of peak and off-peak conditions: a – performance comparison over trips that happened on a sample day; b – MAPE comparison for various days

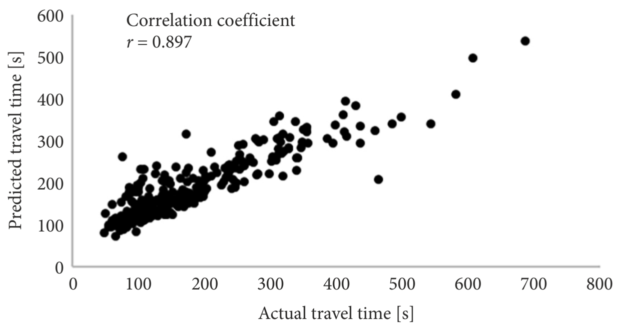


Figure 5. Sample scatter plot of actual vs. predicted travel times (temporal SVM)

variability subsection (47) using temporal SVM. From Figure 5, it can be observed that the predicted values obtained from the temporal SVM are having a strong correlation with the actual travel times with a correlation coefficient of r = 0.897.

A similar analysis has been carried out for all the subsections facing high variability and the results are presented in Table 4. From Table 4, it can be observed that the temporal SVM is performing better than the other methods for the subsections that are having high variability in travel time. From Table 4, it can be observed that the correlation coefficient obtained from temporal SVM is ranging from 0.72 to 0.94 for the sections that were facing high variability, indicating good performance (Cohen 1988; Evans 1995; Russo 2021). In addition, from Table 4, it can be observed that the temporal SVM is performing better than other methods.

In next level, the performance of the proposed spatial and temporal SVM methods were compared with the spatial and temporal KFT across all subsections of

Table 4. Correlation coefficients obtained for various prediction methods for high variability subsections

Subsection index	Temporal SVM	Temporal KFT	Spatial SVM	Spatial KFT
11	0.858	0.475	0.279	0.252
13	0.842	0.548	0.237	0.219
17	0.911	0.717	0.330	0.395
26	0.853	0.728	0.556	0.213
27	0.838	0.606	0.361	0.208
28	0.927	0.536	0.543	0.433
37	0.914	0.551	0.230	0.117
38	0.826	0.533	0.441	0.233
40	0.824	0.525	0.260	0.055
41	0.723	0.514	0.375	0.072
42	0.776	0.639	0.414	0.060
45	0.926	0.560	0.316	0.280
46	0.942	0.520	0.468	0.421
47	0.898	0.519	0.369	0.413

the route for a sample day and the results are presented in Figure 6a. From Figure 6a, it can be observed that the spatial SVM is performing well in low variance sections and temporal SVM is performing better in sections that are having high variance (subsection ID's as shown in Table 2) within a MAPE of 20%. According to Lewis' scale of interpretation of prediction accuracy (Lewis 1982), any forecast with a MAPE value <10% can be considered highly accurate, 11…20% as good, 21…50% as reasonable and >51% as inaccurate. According to this, the prediction accuracy of the proposed method can be considered as good.

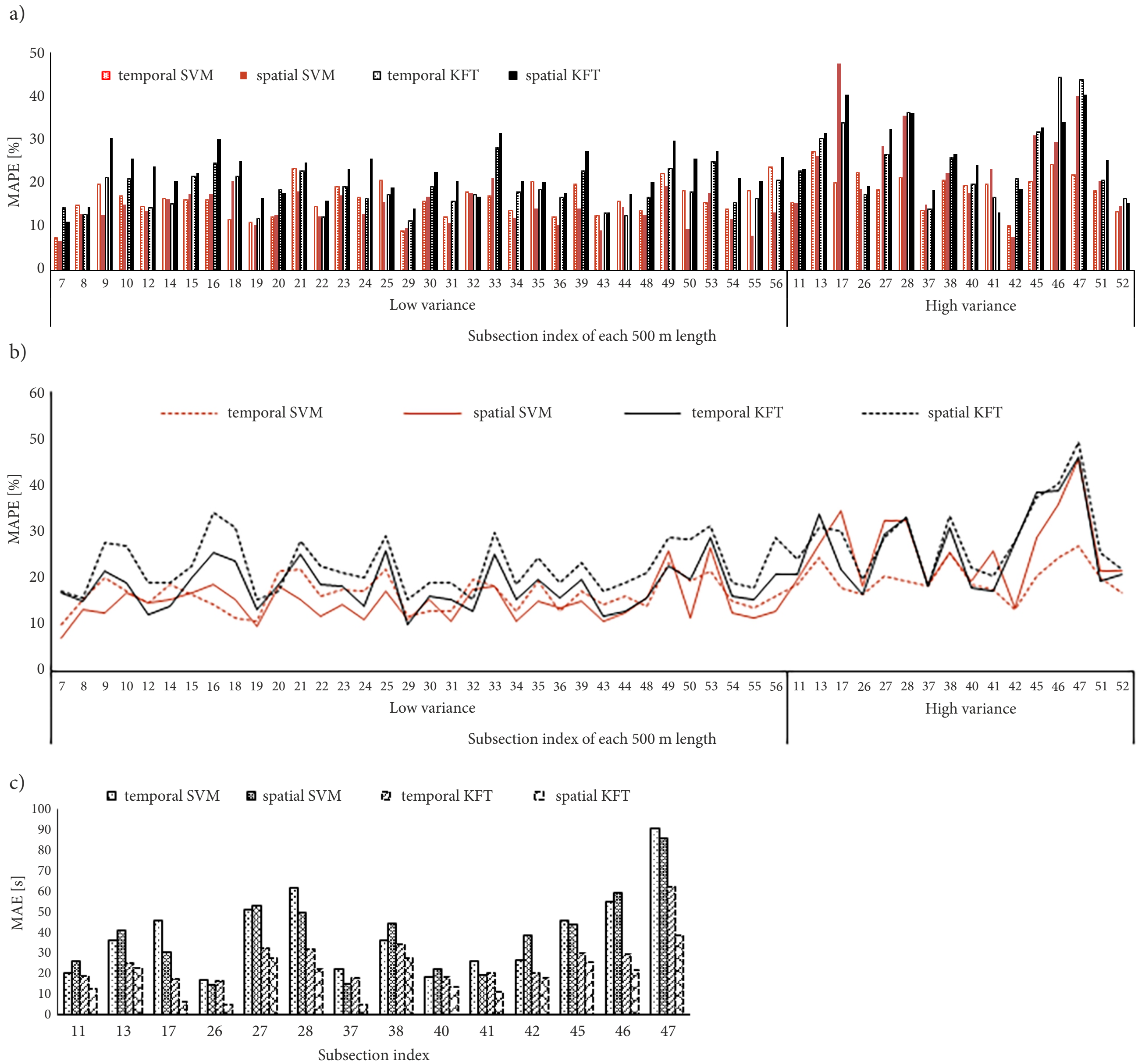


Figure 6. Performance comparison over subsections: a – for all methods in a sample day; b – for SVM methods; c – comparison of MAE for all prediction methods

Performance comparison across multiple days was also carried out and Figure 6b shows the results in terms of MAPE for all sections along the route. It can be observed that temporal SVM outperforms all the other approaches in high variance sections and spatial SVM outperforms other approaches in low variance sections too.

Since, the user feels the error in terms of actual deviations, a performance comparison was also made over different subsections of 500 m length along the considered routes. Figure 6c shows the comparison of average deviation of predicted travel time for all high variability sections along the route (worst case sections) using both spatial and temporal models using KFT and SVM. From Figure 6c, it can be observed that the deviations obtained from temporal SVM were much lesser than the other methods (spatial SVM, spatial KFT and temporal KFT).

From the above results, it can be seen that SVM is performing better than KFT based methods and hence KFT results are not included in further analysis. Among the SVM approaches, temporal SVM was able to perform best under high variability conditions and was able to make better predictions. For the sections that were having lower variability, spatial SVM has slight advantage over temporal SVM methods. In addition, the data requirements are much lesser for spatial SVM since a temporal database is not required in that case. Hence, it would be ideal to use a combination of spatial and temporal SVM, with temporal being used for high variance sections alone and spatial SVM for all the other sections. However, this required a methodology to choose between spatial and temporal SVM for each subsection.

For this, the thresholds were identified based on the mean and variance of the travel times of the subsection of interest. With these thresholds, a subsection was classified as LMLV subsection or HMHV subsection. Figure 7a and Figure 7b show variation in MAPE with increasing mean and standard deviation of travel time. From Figure 7a and Figure 7b, it can be observed that spatial SVM model is performing better than temporal SVM model if a sections' mean travel time is less than 100 s and standard deviation is less than 65 s. In other words, temporal models are performing better if the mean travel time of the subsection is more than 100 s and standard deviation is more than 65 s and hence these values were chosen as the threshold values.

Using the above identified thresholds, all the sections were classified into LMLV and HMHV so that suitable models can be applied to predict bus travel time. Figure 8 shows all the sections classified to LMLV and HMHV sections using the above identified thresholds.

From Figure 8, it can be seen that most of the HMHV sections are within the city area (trip starts in a suburban area and ends in the city area). However, depending on the traffic or other environmental conditions, a subsection, which is found to be LMLV may become HMHV and vice versa. For example, sections that fall under LMLV may become HMHV during peak hours. Thus, it will be more meaningful to classify these sections into LMLV or HMHV dynamically based on real-time conditions than the offline method discussed earlier. Therefore, an automated procedure was developed to select either of spatial SVM or temporal SVM, based on current travel times and is discussed next.

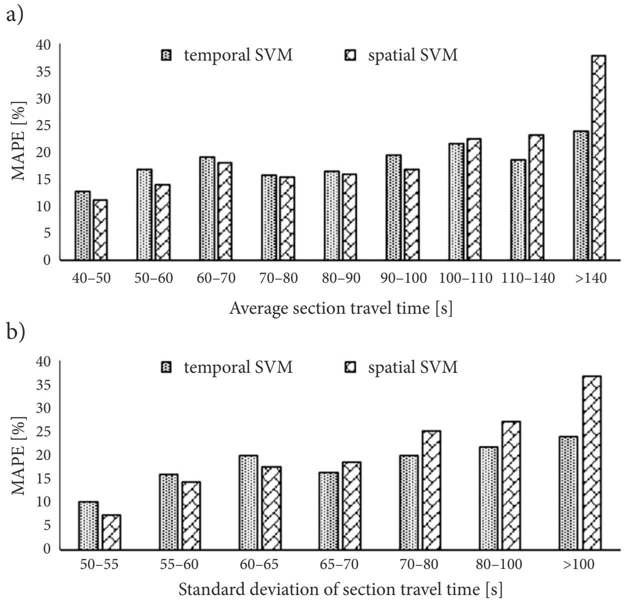


Figure 7. Variation in MAPE with increase in: a – mean travel time; b – standard deviation of travel time

5. Automated systems performance

The above discussed methodology to identify HMHV and LMLV was automated by dynamically classifying each section under consideration to HMHV or LMLV by comparing the latest travel time obtained from each section with the identified threshold mean. Thus, no particular section will be fixed as LMLV or HMHV and vary dynamically depending on the latest data obtained from the field. Therefore, while implementing the prediction algorithm in real-time, the same section may use temporal or spatial SVM models based on recently reported travel times from that section. The optimum number of previous trips to be used for this was chosen heuristically corresponding to minimum MAPE. From the analysis, it was observed that the least MAPE was obtained when 3 previous trips travel times were used as input. Hence, the previous 3 trips were considered for calculating the mean and standard deviation, based on which choice between spatial or temporal SVM models would be made. Thus, if the mean of the 3 previous trips exceed 100 s, the temporal SVM models are used to predict the travel time and spatial models are used otherwise.

To illustrate the performance of the dynamic selection of prediction method, a sample rainy day was selected, in which the travel time variation was found to be very high as shown in Figure 9. Figure 9 shows the change in mean and standard deviation of travel time along different sections during rainy days. On comparing Figure 9 with Figure 8, it can be observed that subsections such as 48, 50, 54, and 55 became HMHV, which were originally LMLV.

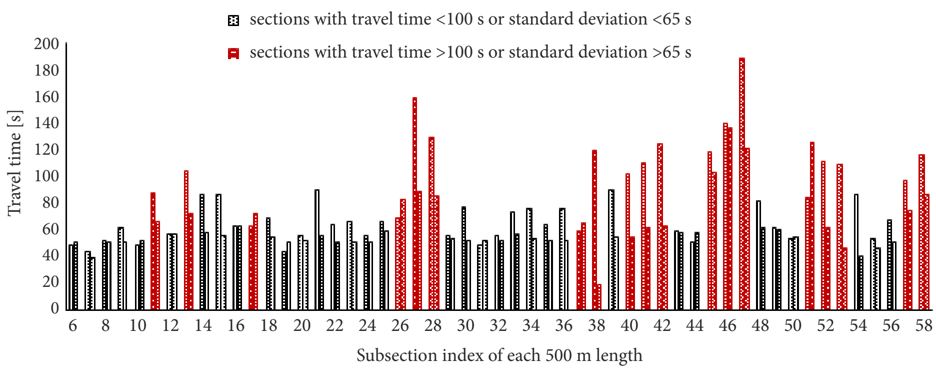


Figure 8. HMHV and LMLV sections identified based on fixed thresholds

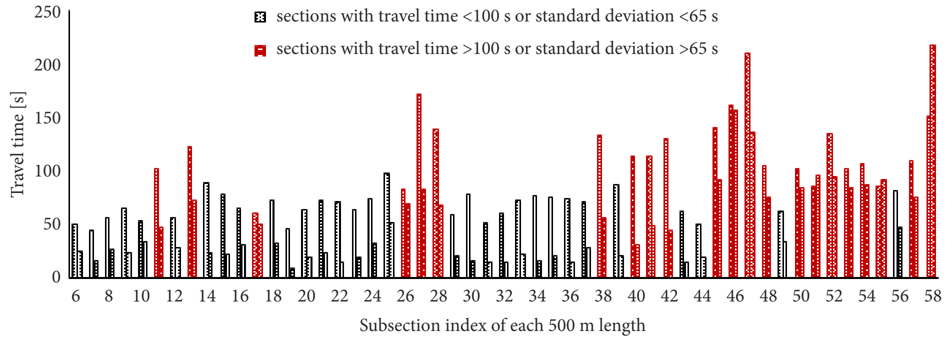


Figure 9. Variation of mean and std. deviation of travel time for various sections in rainy days

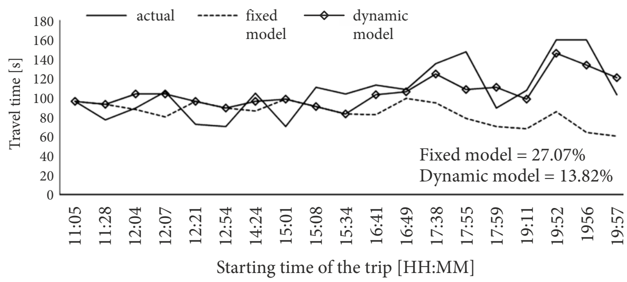


Figure 10. Performance evaluation of a sample subsection (54) on a rainy day

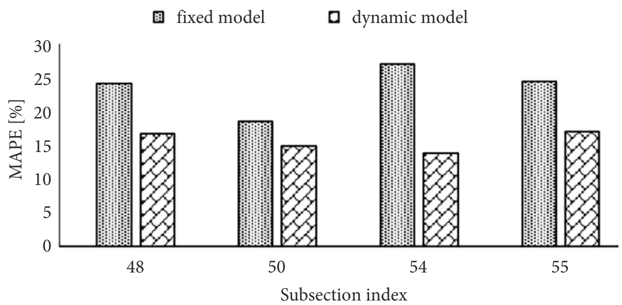


Figure 11. MAPE comparison between fixed and dynamic models in rainy days

A detailed analysis of one of these sections, which changed from LMLV to HMHV, was carried out and results obtained are presented in Figure 10. Figure 10 shows a sample comparison of the predicted and measured travel times for a sample HMHV subsection using both fixed and dynamic models. It can be observed that dynamic model is able to capture the variations in this subsection better than the fixed model with a MAPE reduction from 27 to 13%. Results obtained for all such sections that changed between HMHV or LMLV are shown in Figure 11.

From Figure 11, it can be observed that dynamic model is able to capture the change in travel time better than fixed model, showing the advantage of automating the process.

Summary and conclusions

The heterogeneity and lack of lane discipline makes the Indian traffic highly varying and hence most of the existing solutions that were developed for homogeneous and lane disciplined traffic conditions may not work under such conditions. The present study was an attempt for developing a real-time bus arrival prediction system paying special attention to this high variance. Analysis was carried out using the GPS data collected from buses running in route number 19B in Chennai (India). The main highlights and contributions of the present study are:

- the present study developed a real-time bus arrival prediction system that can capture the variations in the system using SVM. This is one of the first attempts to use SVM for bus arrival time prediction for a traffic system with very high variance such as the heterogeneous and lane-less Indian traffic;
- one of the unique features of the present study in terms of the traffic system under consideration and the variables used as input for prediction. It used only the travel time collected using GPS units fitted in buses as inputs. Another difference from earlier studies is in terms of input quantity selection. The optimum amount of data required to predict the next trip travel time was found using ApEn technique;
- two models were developed, namely spatial SVM and temporal SVM, to predict bus travel time. It was observed that in high mean and variance sections, temporal models are performing better than spatial. An algorithm to dynamically choose between the spatial and temporal SVM models, based on the current travel time, was also developed. Results showed that this automated switching was able to capture the variations better and produce accurate results making it feasible for real-time field implementations.

Overall, the present study focused on accurate travel time prediction by focusing more on sections that are facing high variability. The results showed that the automated system performance of the proposed prediction model was able to capture the variations in traffic conditions better than existing methods and produce accurate results making it feasible for field applications.

Funding

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

References

Abdelfattah, A. M.; Khan, A. M. 1998. Models for predicting bus delays, *Transportation Research Record: Journal of the Transportation Research Board* 1623: 8–15. https://doi.org/10.3141/1623-02

Bian, B.; Zhu, N.; Ling, S.; Ma, S. 2015. Bus service time estimation model for a curbside bus stop, *Transportation Research Part C: Emerging Technologies* 57: 103–121. https://doi.org/10.1016/j.trc.2015.06.011

Cathey, F. W.; Dailey, D. J. 2003. A prescription for transit arrival/departure prediction using automatic vehicle location data, *Transportation Research Part C: Emerging Technologies* 11(3–4): 241–264. https://doi.org/10.1016/s0968-090x(03)00023-8

Chamberlain, R. G. 1996. *GIS FAQ Q5.1: Great Circle Distance between 2 Points*. Available from Internet: https://www.movable-type.co.uk/scripts/gis-faq-5.1.html

Chang, C.-C.; Lin, C.-J. 2011. LIBSVM: a library for support vector machines, *ACM Transactions on Intelligent Systems and Technology* 2(3): 27. https://doi.org/10.1145/1961189.1961199

Chen, M.; Liu, X.; Xia, J.; Chien, S. I. 2004. A dynamic bus-arrival time prediction model based on APC data, *Computer-Aided Civil and Infrastructure Engineering* 19(5): 364–376. https://doi.org/10.1111/j.1467-8667.2004.00363.x

Chen, S.-K.; Zhan, C.-C.; Chen, L.-G. 2007. Prediction method of bus arrival time based on link travel time, *Computer Engineering* 33(21): 281–282. (in Chinese). https://doi.org/10.3969/j.issn.1000-3428.2007.21.100

Cheng, S.; Liu, B.; Zhai, B. 2010. Bus arrival time prediction model based on APC data, in *6th Advanced Forum on Transportation of China* (*AFTC 2010*), 16 October 2010, Beijing, China, 165–169. https://doi.org/10.1049/cp.2010.1123

Chu, L.; Oh, J.-S.; Recker, W. 2005. Adaptive Kalman filter based freeway travel time estimation, in *Transportation Research Board 84th Annual Meeting Compendium of Papers CD-ROM*, 9–13 January 2005, Washington, DC, US, 1–21.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Routledge. 400 p.

Dailey, D. J.; Maclean, S. D.; Cathey, F. W.; Wall, Z. R. 2001. Transit vehicle arrival prediction: algorithm and large-scale implementation, *Transportation Research Record: Journal of the Transportation Research Board* 1771: 46–51. https://doi.org/10.3141/1771-06

Evans, J. D. 1995. *Straightforward Statistics for the Behavioral Sciences*. Brooks/Cole Pub Co. 634 p.

Fan, W.; Gurmu, Z. 2015. Dynamic travel time prediction models for buses using only GPS data, *International Journal of Transportation Science and Technology* 4(4): 353–366. https://doi.org/10.1016/S2046-0430(16)30168-X

Fatima, E.; Kumar, R. 2014. Introduction of public bus transit in Indian cities, *International Journal of Sustainable Built Environment* 3(1): 27–34. https://doi.org/10.1016/j.ijsbe.2014.06.001

Fu, L.; Rilett, L. R. 1998. Expected shortest paths in dynamic and stochastic traffic networks, *Transportation Research Part B: Methodological* 32(7): 499–516. https://doi.org/10.1016/s0191-2615(98)00016-2

Hans, E.; Chiabaut, N.; Leclercq, L.; Bertini, R. L. 2015. Real-time bus route state forecasting using particle filter and mesoscopic modeling, *Transportation Research Part C: Emerging Technologies* 61: 121–140. https://doi.org/10.1016/j.trc.2015.10.017

Jeong, R.; Rilett, L. R. 2004. Bus arrival time prediction using artificial neural network model, in *Proceedings: the 7th International IEEE Conference on Intelligent Transportation Systems*, 3–6 October 2004, Washington, DC, US, 988–993. https://doi.org/10.1109/itsc.2004.1399041

Kumar, S. V.; Dogiparthi K. C.; Vanajakshi, L.; Subramanian, S. C. 2017. Integration of exponential smoothing with state space formulation for bus travel time and arrival time prediction, *Transport* 32(4): 358–367. https://doi.org/10.3846/16484142.2015.1100676

Kumar, S. V.; Vanajakshi, L. 2012. Application of multiplicative decomposition and exponential smoothing techniques for bus arrival time prediction, in *TRB 91st Annual Meeting Compendium of Papers DVD*, 22–26 January 2012, Washington, DC, US, 1–12.

Kumar, S. V.; Vanajakshi, L. 2014. Pattern identification based bus arrival time prediction, *Proceedings of the Institution of Civil Engineers – Transport* 167(3): 194–203. https://doi.org/10.1680/tran.12.00001

Kumar, B. A.; Vanajakshi, L.; Subramanian, S. C. 2014a. Pattern-based bus travel time prediction under heterogeneous traffic conditions, in *TRB 93rd Annual Meeting Compendium of Papers*, 12–16 January 2014, Washington, DC, US, 1–16.

Kumar, V.; Kumar, B. A.; Vanajakshi, L.; Subramanian, S. C. 2014b. Comparison of model based and machine learning approaches for bus arrival time prediction, in *TRB 93rd Annual Meeting Compendium of Papers*, 12–16 January 2014, Washington, DC, US, 1–14.

Kwon, J.; Coifman, B.; Bickel, P. 2000. Day-to-day travel-time trends and travel-time prediction from loop-detector data, *Transportation Research Record: Journal of the Transportation Research Board* 1717: 120–129. https://doi.org/10.3141/1717-15

Lewis, C. D. 1982. *Industrial and Business Forecasting Methods: a Practical Guide to Exponential Smoothing and Curve Fitting*. Butterworth Scientific. 143 p.

Li, R. 2006. *Enhancing Motorway Travel Time Prediction Models through Explicit Incorporation of Travel Time Variability*. PhD Thesis. Monash University, Melbourne, Australia. 264 p.

Liu, H.; Van Zuylen, H. J.; Van Lint, H.; Chen, Y.; Zhang, K. 2005. Prediction of urban travel times with intersection delays, in *Proceedings: 2005 IEEE Intelligent Transportation Systems*, 13–16 September 2005, Vienna, Austria, 402–407. https://doi.org/10.1109/ITSC.2005.1520198

Liu, X.; Chien, S. I.; Chen, M. 2014. An adaptive model for highway travel time prediction, *Journal of Advanced Transportation* 48(6): 642–654. https://doi.org/10.1002/atr.1216

Makridakis, S. 1993. Accuracy measures: theoretical and practical concerns, *International Journal of Forecasting* 9(4): 527–529. https://doi.org/10.1016/0169-2070(93)90079-3

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: Technology, Planning, and Operations* 15(2): 75–90. https://doi.org/10.1080/15472450.2011.570109

Nanthawichit, C.; Nakatsuji, T.; Suzuki, H. 2003. Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway, *Transportation Research Record: Journal of the Transportation Research Board* 1855: 49–59. https://doi.org/10.3141/1855-06

Padmanaban, R. P. S.; Divakar, K.; Vanajakshi, L.; Subramanian, S. C. 2010. Development of a real-time bus arrival prediction system for Indian traffic conditions, *IET Intelligent Transport Systems* 4(3): 189–200. https://doi.org/10.1049/iet-its.2009.0079

Patnaik, J.; Chien, S.; Bladikas, A. 2004. Estimation of bus arrival times using APC data, *Journal of Public Transportation* 7(1): 1–20. https://doi.org/10.5038/2375-0901.7.1.1

Pattanamekar, P.; Park, D.; Rilett, L. R.; Lee, J.; Lee, C. 2003. Dynamic and stochastic shortest path in transportation networks with two components of travel time uncertainty, *Transportation Research Part C: Emerging Technologies* 11(5): 331–354. https://doi.org/10.1016/s0968-090x(03)00029-9

Pincus, S. M. 1991. Approximate entropy as a measure of system complexity, *Proceedings of the National Academy of Sciences of the United States of America* 88(6): 2297–2301. https://doi.org/10.1073/pnas.88.6.2297

Rajbhandari, R. 2005. *Bus Arrival Time Prediction Using Stochastic Time Series and Markov Chains*. PhD Dissertation. New Jersey Institute of Technology, Newark, NJ, US. 157 p. Available from Internet: http://archives.njit.edu/vol01/etd/2000s/2005/njit-etd2005-038/njit-etd2005-038.pdf

Ramakrishna, Y.; Ramakrishna, P.; Lakshmanan, V.; Sivanandan, R. 2006. Bus travel time prediction using GPS data, in *Map India 2006: 9th Annual International Conference and Exhibition*, 30 January – 1 February 2006, Delhi, India.

Russo, R. 2021. *Statistics for the Behavioural Sciences: an Introduction to Frequentist and Bayesian Approaches*. Routledge. 330 p.

Schölkopf, B.; Smola, A. J.; Williamson, R. C.; Bartlett, P. L. 2006. New support vector algorithms, *Neural Computation* 12(5): 1207–1245. https://doi.org/10.1162/089976600300015565

Shalaby, A.; Farhan, A. 2004. Prediction model of bus arrival and departure times using AVL and APC data, *Journal of Public Transportation* 7(1): 41–61. https://doi.org/10.5038/2375-0901.7.1.3

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

Vanajakshi, L.; Rilett, L. R. 2004. A comparison of the performance of artificial neural networks and support vector machines for the prediction of traffic speed, in *IEEE Intelligent Vehicles Symposium, 2004*, 14–17 June 2004, Parma, Italy, 194–199. https://doi.org/10.1109/ivs.2004.1336380

Vanajakshi, L.; Rilett, L. R. 2007. Support vector machine technique for the short term prediction of travel time, *2007 IEEE Intelligent Vehicles Symposium*, 13–15 June 2007, Istanbul, Turkey, 600–605. https://doi.org/10.1109/ivs.2007.4290181

Vanajakshi, L.; Subramanian, S. C.; Sivanandan, R. 2009. Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses, *IET Intelligent Transport Systems* 3(1): 1–9. https://doi.org/10.1049/iet-its:20080013

Vapnik, V. N. 1999. An overview of statistical learning theory, *IEEE Transactions on Neural Networks* 10(5): 988–999. https://doi.org/10.1109/72.788640

Wall, Z.; Dailey, D. J. 1999. An algorithm for predicting the arrival time of mass transit vehicles using automatic vehicle location data, in *Transportation Research Board 78th Annual Meeting*, 10–14 January 1999, Washington DC, US, 1–10.

Wu, C.-H.; Su, D.-C.; Chang, J.; Wei, C.-C.; Ho, J.-M.; Lin, K.-J.; Lee, D. T. 2003. An advanced traveler information system with emerging network technologies, in *Proceedings of 6th Asia-Pacific Intelligent Transportation Systems Forum*, 4–9 October 2003, Taipei, Taiwan, 1–8.

Xu, H., Ying, J. 2017. Bus arrival time prediction with real-time and historic data, *Cluster Computing* 20(4): 3099–3106. https://doi.org/10.1007/s10586-017-1006-1

Yang, M.; Chen, C.; Wang, L.; Yan, X.; Zhou, L. 2016. Bus arrival time prediction using support vector machine with genetic algorithm, *Neural Network World* 26(3): 205–217. https://doi.org/10.14311/nnw.2016.26.011

Yu, B.; Lam, W. H. K.; Tam, M. L. 2011. Bus arrival time prediction at bus stop with multiple routes, *Transportation Research Part C: Emerging Technologies* 19(6): 1157–1170. https://doi.org/10.1016/j.trc.2011.01.003

Yu, B.; Song, X.; Guan, F.; Yang, Z.; Yao, B. 2016. *k*-nearest neighbor model for multiple-time-step prediction of short-term traffic condition, *Journal of Transportation Engineering* 142(6): 04016018. https://doi.org/10.1061/(asce)te.1943-5436.0000816

Yu, B.; Yang, Z.; Yao, B. 2006. Bus arrival time prediction using support vector machines, *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations* 10(4): 151–158. https://doi.org/10.1080/15472450600981009

Yu, Bin; Yang, Z.-Z.; Chen, K.; Yu, Bo. 2010a. Hybrid model for prediction of bus arrival times at next station, *Journal of Advanced Transportation* 44(3): 193–204. https://doi.org/10.1002/atr.136

Yu, Bo; Lu, J.; Yu, Bin; Yang, Z. 2010b. An adaptive bus arrival time prediction model, *Journal of the Eastern Asia Society for Transportation Studies* 8: 1126–1136.

Yu, Z.; Wood, J. S; Gayah, V. V. 2017. Using survival models to estimate bus travel times and associated uncertainties, *Transportation Research Part C: Emerging Technologies* 74: 366–382. https://doi.org/10.1016/j.trc.2016.11.013

Zhou, Y.; Yao, L.; Chen, Y.; Gong, Y.; Lai, J. 2017. Bus arrival time calculation model based on smart card data, *Transportation Research Part C: Emerging Technologies* 74: 81–96. https://doi.org/10.1016/j.trc.2016.11.014