



A Probabilistic Travel Time Modeling Approach Based on Spatiotemporal Speed Variations

Mohammed Elhenawy¹, Abdallah A. Hassan²,
and Hesham A. Rakha^{2,3,4}✉

¹ Centre for Accident Research and Road Safety,
University of Technology at the Queensland, Queensland, Australia
elhenawy@vt.edu

² Bradley Department of Electrical and Computer Engineering,
Electrical and Computer Engineering, MC 0111, 1185 Perry St.,
Blacksburg 24061, VA, USA
{abdallah, hrakha}@vt.edu

³ Charles E. via, Jr. Department of Civil and Environmental Engineering,
Blacksburg, VA 24061, USA

⁴ Center for Sustainable Mobility, Virginia Tech Transportation Institute,
Blacksburg, VA 24061, USA

Abstract. The rapid development and deployment of Intelligent Transportation Systems (ITSs) require the development of data driven algorithms. Travel time modeling is an integral component of travel and transportation management and travel demand management functions. Travel time has a massive impact on driver's route choice behavior and the assessment of the transportation system performance. In this paper, a mixture of linear regression is proposed to model travel times. The mixture of linear regression models has three advantages. First, it provides better model fitting compared to simple linear regression. Second, the proposed model can capture the bi-modal nature of travel time distributions and link it to the uncongested and congested traffic regimes. Third, the means of the bi-modal distributions are modeled as functions of the input predictors. This last advantage allows for the quantitative evaluation of the probability of each travel time state as well as the uncertainty associated with each state at any time of the day given the values of the predictors at that time. The proposed model is applied to archived data along a 74.4-mile freeway stretch of I-66 eastbound to connect I-81 and Washington D.C. The experimental results show the ability of the model to capture the stochastic nature of the travel time and gives good travel time predictions.

Keywords: Travel time modelling · Travel time reliability · Mixture of linear regressions

1 Introduction

Minimizing drivers' travel times from their origins to their destinations is a major Intelligent Transportation Systems (ITSs) objective. However, it is also extremely challenging due to the dynamic nature of traffic flow, which is, in most cases, highly unpredictable. One straightforward strategy involves directing vehicles or guiding drivers to follow routes that avoid congested paths. A critical step for this route planning or guidance to be effective is the ability to accurately predict travel times of different alternative routes from source to destination.

In addition, travel time represents an important performance measure for traffic system evaluation. It is easily understood by drivers and operators of traffic management systems, and can be viewed as a simple summary of a traffic system's complex behavior. In order for an ITS to accurately predict the travel time, it must have the following capabilities, each of which comes with associated difficulties:

1. Sensing and acquiring the current state of the transportation network of interest where a number of data values need to be detected and collected, including traffic conditions and parameters at different parts of the network, whether some roads are currently congested, current weather conditions, time of day, whether there is an incident on any road in the network, etc. Gathering such data on every road and intersection with the quality that allows accurate forecasting of travel time between two points in the network may be fairly expensive.
2. Storing a long history of traffic parameters for the transportation network of interest to support future prediction of travel times. This historical dataset may be large and difficult to use and manage.
3. Feeding the current state of the network along with its traffic history to some type of model that predicts travel time if a trip will start from some point and end in another in the network at some specific time. Designing such a model is challenging, as is finding a set of current or historical parameters with real prediction power. The most useful model may be road dependent, and even for a single road, it has been shown that different models may describe the traffic behavior more accurately at different traffic conditions. For instance, one model may be more useful when the road is congested, while another model may be more accurate when vehicles are flowing freely, etc.

In short, accurate traffic time prediction is challenging due to the high cost of sensing and collecting enough useful current and historical traffic data. Even when such data is available, it is still difficult to determine which type of model best describes the traffic behavior, and which traffic parameters should be fed to the model for the best predictions. Moreover, the best course of action may be to use two or more models and switch between them depending on current traffic conditions. This option adds a new challenge, as it is necessary to decide which model from the set of models will be used for some specific input data, or whether different models will be used for prediction with some weight applied to each output prediction to reach a final travel time prediction.

In this paper, a new method for travel time prediction is proposed. This method uses a mixture of linear regressions motivated by the fact that travel time distribution is not unimodal, since two modes or regimes of traffic can exist—one at congestion state, and the other at free-flow state. We show how the proposed model is very flexible and gives slightly different accuracy when we model the travel time or the log of the travel time using two different set of predictors. The First set of predictors are the selected elements from the spatiotemporal speed matrix based on their estimated importance using random forest. Then these set of speeds will be the input predictors to the statistical model. The second set of predictors are the instantaneous travel time and the average of the historical travel time. The proposed model is built and tested using probe data provided by INRIX and supplemented with traditional road sensor data as well as mobile devices and other sources. The dataset was collected from a freeway stretch of I-66 eastbound connecting I-81 and Washington, D.C. The traffic on this stretch is often extremely heavy, which makes travel time prediction more challenging, but also makes the data more valuable and helps create a more realistic model.

2 Related Work

Various methods and algorithms have been proposed in the literature for travel time prediction. These methods can roughly be classified into two main categories: statistical-based data-driven methods and simulation-based methods. This section focuses on the statistical-based methods since the proposed solution in this paper falls under this class of methods, and because more research in the literature uses statistical methods.

Several researchers fit different regression models to predict travel time. A typical approach is to fit a multiple linear regression (MLR) model using explanatory variables representing instantaneous traffic state and historical traffic data, as, for example, [1, 2]. The model proposed in [1] was even able to use a single linear regression (SLR) to successfully provide acceptable travel time predictions. Some researchers developed hybrid methods where a regression model was used in conjunction with other advanced statistical methods. For example, [3] used regression with statistical tree methods. Another approach [4] proposed an SLR model using bus travel time to predict automobile travel time.

Regression models are generally powerful in predicting travel time for short-term prediction, whereas long-term predictions are less accurate. Regression models are also reported to be more suitable for use in free-flow rather than congested traffic, and fail to accurately predict when incidents have occurred [5].

The idea of using a mixture models for different traffic regimes has also previously been explored [6]. The model developed in this paper attempts to overcome the drawbacks of previous work that used mixture models of two or three components to model travel time reliability, which suffer from the following limitations:

1. The mean of each component is not modeled as a function of the available predictors.
2. The proportion variable is fixed at each time slot, which limits the model's flexibility.

- Information provided given the time slot of the day is the probability of each component (fixed) and the 90th percentile.

Another class of statistical-based methods in literature uses time series models for travel time prediction, using, for example, auto-regressive prediction models [7–9], multivariate time series models [10], and the auto-regressive integrated moving average (ARIMA) technique [11]. Similar to regression models, time series models are more suitable for free-flow traffic than for congested traffic, may fail with unusual incidents, and are more accurate for short-term predictions [5].

Another common technique used for travel time prediction is the use of artificial neural networks. A feed-forward neural network is used in [12] to predict journey time. Later, more advanced neural network techniques were used to model and predict travel time [13–19]. Accurate predictions were achieved for most proposed models; for example, in [20] the prediction error was only 4%.

3 Methods

The definitions of historical, instantaneous and ground truth travel times are introduced in this section. In addition, we present a brief introduction of the powerful modelling technique used in this paper. Expectation-maximization (EM) is used to fit the mixture of linear regression models to the historical data.

3.1 Travel Time Ground Truth Calculation

The calculation of the travel time ground truth is based on trajectory construction and the known speed through the trajectory's cells. A simple example of travel time ground truth calculation based on trajectory construction is demonstrated in Fig. 1. In this example, the roadway is divided into four sections using segments of length Δx and a time interval of Δt . We assumed that the speed is homogenous within each cell. The average speed of the red-dotted cell ($i = 2, n = 3$) in the figure is $u(x2, t3)$. Consequently, the trajectory slope represents the speed in each cell. Once the vehicle enters a new cell, the trajectory within this cell can be drawn as the straight blue line in Fig. 1 using the cell speed as the slope. Finally, the ground truth travel time can be calculated when the trip reaches the downstream boundary of the last freeway section. It should be noted that the ground truth travel times were computed using the same dataset and used as the response (y).

3.2 Instantaneous Travel Time

The instantaneous method is very simple where it assumes the segment speed does not change during the entire trip time. The travel time using the instantaneous approach is shown in Eq. (1).

$$\text{instantaneous travel time} = \sum_{i=1}^h \frac{L_i}{v_i^{t0}} \quad (1)$$

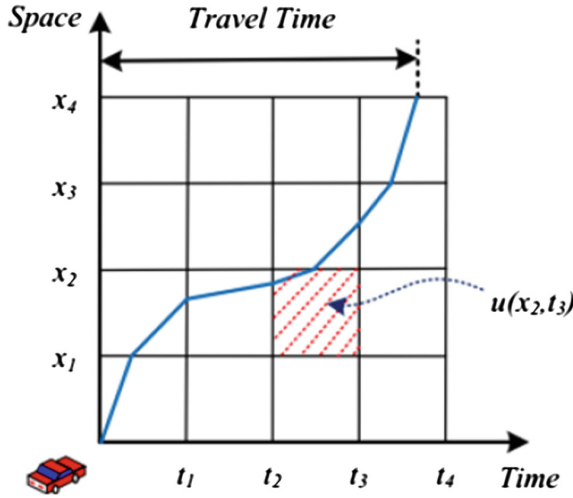


Fig. 1. Illustration of travel time ground truth calculation.

Where

L_i is the length of segment i

v_i^{t0} is the speed at segment i at the departure time $t0$

h is the total number of segments.

3.3 Historical Average Method

If the spatiotemporal speed matrices are known for several previous months, then the ground truth travel time at each time interval for each day can be calculated. The historical average at any time of day D is calculated using Eq. (2).

$$\text{average historical travel time} = \sum_{i=1}^{Z_{D_i}} \frac{GTTT_{D_i}^{t0}}{Z_{D_i}} \forall D_i = \text{Saturday}, \dots, \text{Friday} \quad (2)$$

Where $GTTT_{D_i}^{t0}$ is the ground truth travel time at departure time $t0$ at historical day D_i and Z_{D_i} is number of days included in the average.

3.4 Mixture of Linear Regressions

A mixture of linear regressions was studied carefully [21, 22]. It can be used to model travel time under different traffic regimes. The mixture of linear regression can be written as:

$$f(y|X) = \sum_{j=1}^m \frac{\lambda_j}{\sigma_j \sqrt{2\pi}} e^{-\frac{(y-x^T \beta_j)^2}{2\sigma_j^2}} \quad (3)$$

where y_i is the response corresponding to a vector p of predictors; x_i^T, β_j is the vector of regression coefficients for the j^{th} component and λ_j is mixing probability of the j^{th} component.

The model parameters $\psi = \{\beta_1, \beta_2, \dots, \beta_m, \sigma_1^2, \sigma_2^2, \dots, \sigma_m^2, \lambda_1, \lambda_2, \dots, \lambda_m\}$. can be estimated by maximizing the log-likelihood of Eq. (1) given a set of response predictor pairs $(y_1, x_1), (y_2, x_2), \dots, (y_n, x_n)$ using an EM algorithm. The EM algorithm iteratively finds the maximum likelihood estimates by alternating the E-step and M-step. Let $\psi^{(k)}$ be the parameters' estimates after the k^{th} iteration. In the E-step, the posterior probability of the i^{th} observation from component j is computed using Eq. (4).

$$w_{ij}^{(k+1)} = \frac{\lambda_j^{(k)} \phi_j(y_i|x_i, \psi^{(k)})}{\sum_{j=1}^m \lambda_j^{(k)} \phi_j(y_i|x_i, \psi^{(k)})} \tag{4}$$

where $\phi_j(y_i|x_i, \psi^{(k)})$ is the probability density function of the j^{th} component.

In the M-step, the new parameters' estimates $\psi^{(k+1)}$ that maximize the log-likelihood function in Eq. (3) are calculated using Eqs. (5–7)

$$\lambda_j^{(k+1)} = \frac{\sum_{i=1}^n w_{ij}^{(k+1)}}{n} \tag{5}$$

$$\hat{\beta}_j^{(k+1)} = (X^T W_j X)^{-1} X^T W_j \tag{6}$$

where X is the predictors' matrix with n rows and $(p + 1)$ columns, Y is the corresponding $n \times 1$ response vector, and W is an $n \times n$ diagonal matrix which has $w_{ij}^{(k+1)}$ on its diagonal.

$$\hat{\sigma}_j^{2(k+1)} = \frac{\sum_{i=1}^n w_{ij}^{(k+1)} (y_i - x_i^T \hat{\beta}_j^{(k+1)})^2}{\sum_{i=1}^n w_{ij}^{(k+1)}} \tag{7}$$

The E-step and M-step are alternated repeatedly until the change in the incomplete log-likelihood is arbitrarily small as shown in Eq. (8).

$$\left| \prod_{i=1}^n \sum_{j=1}^m \lambda_j^{(k+1)} \phi_j(y_i|x_i, \psi^{(k+1)}) - \prod_{i=1}^n \sum_{j=1}^m \lambda_j^{(k)} \phi_j(y_i|x_i, \psi^{(k)}) \right| < \xi \tag{8}$$

where ξ is a small number.

4 The Predictors' Sets

In this section, we describe the two approaches used to form the predictors' sets. In the first approach, the random forest machine-learning algorithm (RF) is used to select a subset of important predictors for travel time modelling. Where the second approach incorporate information from the past by using the historical travel time and summarize the current speeds within a window starting right before the departure time t_0 using the instantaneous travel time.

4.1 The First Set of Predictors

The I-66 stretch of the freeway section used for this research consists of 64 segments. The dataset comprises the spatiotemporal speed matrices for every day in 2013. The default approach for modelling and predicting travel time was to take all the speeds within a window starting right before the departure time t_0 and covering L past time slots back to time $t_0 - L$. Setting $L = 30$ min for example, the number of predictors will be $64 * 6$ at 5 min time aggregation. In order to reduce the dimensions of the predictors' vector, RF is used to select the most important predictors for the travel time model. Steps to select the most important predictors are as follows [23]:

1. For each month, build an RF consisting of 100 trees and find the out-of-bag samples that are not used in the training for each tree.
2. Find the mean square error MSE_{outofbag} of the RF using the out-of-bag samples.
3. Randomly permute the value for each predictor x_i among the out-of-bag samples and calculate the mean square error $MSE_{\text{outofbag}}^{\text{permuted}x_i}$ of the RF.
4. Finally, rank the predictors in descending order based on the $\frac{1}{12} \sum_{\text{month}=1}^{12} (MSE_{\text{outofbag}}^{\text{permuted}x_i} - MSE_{\text{outofbag}})$ and choose the top m ranked predictors.

The higher the predictor's rank in step 4, the more important that predictor. The ranking result shows that, most of the important predictors are speeds of recent segments ($t_0 - 5$). In addition to speed predictors chosen by RF, the historical average travel time at t_0 given the day of the week is added as a predictor.

4.2 The Second Set of Predictors

The other set of predictors are the instantaneous travel time and the average of historical travel time. For example, if we are interested in the travel time reliability at t_0 on day D , the predictor vector will be the instantaneous travel time at the times $\{t_0 - 45, t_0 - 40, \dots, t_0 - 5\}$ and the average of the historical travel times on days D at times $\{t_0, t_0 + 5, \dots, t_0 + 45\}$. Figure 2 shows the average of the historical travel time for each day of the week. There are two peaks of the travel time during morning and evening hours. The height of the peaks is different from one day to another especially between weekdays and weekends.

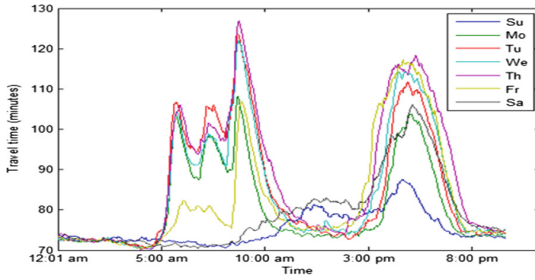


Fig. 2. The average of historical travel time for each day of the week.

5 Data Description

The freeway stretch of I-66 eastbound connecting I-81 and Washington, D.C. was selected as the test site for this study. High traffic volumes are usually observed during morning and afternoon peak hours on I-66 heading towards Washington, D.C., making it an excellent environment to test travel time models.

The traffic data was provided by INRIX, which mainly collects probe data by GPS-equipped vehicles, supplemented with traditional road sensor data, along with mobile devices and other sources [24]. The probe data covers 64 freeway segments with a total length of 74.4 miles. The average segment length is 1.16 miles, and the length of each segment is unevenly divided in the raw data from 0.1 to 8.22 miles. Figure 3 shows the study site and deployment of roadway segments. The raw data provides average speed for each roadway segment and was collected at 1-m intervals.

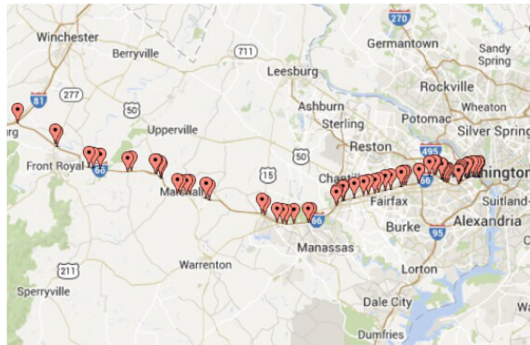


Fig. 3. The study site on I-66 eastbound. (source: Google Maps) [25].

We sorted the raw data was the roadway direction according to each TMC station’s geographic information (e.g., towards eastbound of I-66). Data was examined to check any overlapping or inconsistent stations along the route. Afterward, speed data was aggregated by time intervals (5 min in this study) to reduce noise and smooth measurement errors. This way, the raw data was aggregated to the form of the daily data

matrix along spatial and temporal intervals. Data was missing in the developed data matrix, so data input methods were conducted to estimate the missing data using values of neighboring cells. Finally, the daily spatiotemporal traffic state matrix was generated to model travel time.

6 Experimental Analysis

The experimental work is divided into two subsections. The first subsection is travel time modeling using a mixture of two linear regressions with fixed proportions (λ_1, λ_2) . In this subsection, we will model the travel time and the log of the travel time using the first set and second set of predictors respectively. Consequently, we will show that log-normal model is better than the normal model. The second subsection describes the travel time reliability modelling approach, where we modified the proposed model to allow the proportions to vary as a function of the predictors. The modified model computes the probabilities of encountering free-flow and congested conditions together with the expected and 90-percentile travel times for each regime.

6.1 Modeling Travel Time Using a Mixture of Linear Regressions with Fixed Proportions

The purpose of this section is to experimentally compare the lognormal model and normal model. Each model used a different set of predictor. In other words, the log-normal model will explain the response vector $\log(Y)$ using the corresponding predictors matrix X_2 , where Y is the ground truth travel time and X_2 is the second set of predictors. Where the normal model will explain the response vector Y using the corresponding predictors matrix X_1 . For the sake of completeness, we also compare each model with the corresponding one component linear regression model, which assumes the travel time distribution is uni-modal distribution. To show that, the parameters of the proposed models are estimated using the EM algorithm. Then, two measures are used to compare the models. The Mean Absolute Percentage Error (MAPE) and the Mean Absolute Error (MAE) are used to quantify the errors of both models with respect to the ground truth. MAPE is the average absolute percentage change between the predicted \hat{y}_i^j and the true values y_i^j . MAE is the absolute difference between the predicted and the true values.

$$MAPE = \frac{100}{I \times J} \sum_{j=1}^J \sum_{i=1}^I \frac{|y_i^j - \hat{y}_i^j|}{y_i^j} \tag{9}$$

$$MAE = \frac{1}{I \times J} \sum_{j=1}^J \sum_{i=1}^I |y_i^j - \hat{y}_i^j| \tag{10}$$

Here, J is the total number of days in the testing dataset; I is the total number of time intervals in a single day; and y and \hat{y} denote the ground truth and the predicted

value, respectively, of the travel time for the time interval on the day. The lower the value of these error measures, the better the model.

Table 1 shows values for the MAE and MAPE for normal models using a different number of top-ranked predictors. As shown in Table 1, for all models that are built using a different number of predictors, the models built using the proposed mixture of regressions are better than the linear regression models with smaller MAE and MAPE.

Table 1. Comparison between one and two components normal models using the first set of predictors [25].

p	MAE		MAPE	
	m = 1	m = 2	m = 1	m = 2
6	6.57	5.22	7.19	5.69
11	6.39	5.10	6.99	5.63
16	6.36	5.05	6.96	5.57
21	6.32	5.04	6.89	5.56
26	6.31	5.06	6.90	5.59
31	6.32	5.09	6.90	5.64
36	6.30	5.08	6.88	5.62
41	6.30	5.13	6.88	5.69
46	6.29	5.12	6.87	5.68
51	6.23	5.13	6.80	5.70
56	6.24	5.12	6.82	5.69
61	6.18	5.16	6.77	5.74
66	6.18	5.16	6.76	5.74
71	6.20	5.15	6.79	5.73
76	6.19	5.15	6.78	5.73
81	6.20	5.16	6.79	5.74
86	6.18	5.19	6.78	5.78
91	6.19	5.21	6.79	5.80
96	6.19	5.22	6.79	5.80

Table 2 shows values for the MAE and MAPE for lognormal models using the second set of predictors. Table 2 confirms that two component models are better than one-component models. Moreover, it shows that the lognormal models and the second set of predictors are better than the normal models and the first set of predictors (Table 3).

Different number of predictors are shown in Tables 1 and 2. We tried different number of predictor in order to find the simplest and most accurate model. As shown in the above tables the improvement in the accuracy of the models in terms of MAE and MAPE is not significant. In real time running, we prefer simple models. So that, for the normal model and lognormal model the models which has 16 and 11 predictors respectively are chosen.

Table 2. Comparison between one and two components lognormal models using the second set of predictors.

p	MAE		MAPE	
	m = 1	m = 2	m = 1	m = 2
3	5.04	4.28	5.46	4.58
5	5.04	4.28	5.45	4.58
7	5.03	4.28	5.45	4.58
9	5.03	4.28	5.45	4.58
11	5.03	4.27	5.45	4.57
13	5.03	4.28	5.45	4.58
15	5.03	4.26	5.45	4.56
17	5.03	4.26	5.45	4.56
19	5.03	4.25	5.46	4.56

Table 3. The EM parameters' estimates for a mixture of linear regression assuming normal distribution [25]*.

	1st component	2nd component
<i>intercept</i>	79.4354	96.5943
$x_{29,t0-1}$	-0.0153	-0.0148
$x_{2,t0-1}$	-0.0903	-0.0250
$x_{28,t0-1}$	-0.0668	0.0061
$x_{18,t0-1}$	-0.0912	-0.0519
$x_{27,t0-1}$	0.0187	-0.0449
$x_{40,t0-1}$	-0.2107	-0.1107
$x_{25,t0-1}$	-0.0652	-0.0603
$x_{14,t0-1}$	-0.0245	-0.0136
$x_{29,t0-2}$	-0.0106	-0.0224
$x_{1,t0-1}$	-0.0745	-0.0150
$x_{39,t0-1}$	-0.0174	-0.0331
$x_{21,t0-1}$	-0.0203	-0.0252
$x_{24,t0-1}$	-0.0742	-0.0239
$x_{19,t0-1}$	0.0075	-0.0078
$x_{30,t0-1}$	-0.1269	-0.0558
$x_{13,t0-1}$	0.6767	0.0834
σ_j^2	11.8066	1.7746
λ_j	0.4466	0.5534

*(In this table $x_{(seg\#, time)}$ is the speed at certain segment and time)

Based on the above experimental results we conclude that using the lognormal model is more accurate and simpler in terms of number of predictors than using normal model. So that we can choose the mixture of two linear regression with 11 predictors and use it for the next set of experiments.

Table 4. The EM parameters’ estimates for a mixture of linear regression assuming lognormal distribution.

	1st component	2nd component
<i>intercept</i>	3.4875	3.9663
$x_{(to-25)}^{ins}$	-0.0003	-0.0001
$x_{(to-20)}^{ins}$	-0.0002	0.0001
$x_{(to-15)}^{ins}$	0.0001	-0.0001
$x_{(to-10)}^{ins}$	0.0000	0.0000
$x_{(to-5)}^{ins}$	0.0020	0.0026
$x_{(to)}^{his}$	0.0017	-0.0002
$x_{(to+5)}^{his}$	0.0036	0.0010
$x_{(to+10)}^{his}$	0.0028	0.0012
$x_{(to+15)}^{ins}$	0.0013	-0.0008
$x_{(to+20)}^{his}$	-0.0022	0.0011
$x_{(to+25)}^{ins}$	0.0028	-0.0003
σ_j^2	0.1008	0.0222
λ_j	0.4655	0.5345

6.2 Travel Time Reliability

Travel time reliability is the form of information that we can convey to traveler using the travel time model. Using the proposed model, we can provide the traveler what are the probabilities of congestion and free flow. Moreover, the expected and 90% percentile travel time for each regime can be provided. In order to get good estimates for the above quantities, the proportions should be a function of the predictor which means it varies depending on the values of the predictors. Revisiting the EM algorithm, it estimates the posterior probabilities w_{ij} and the model parameters ψ and returns only ψ at convergence and does not use w_{ij} . As shown in Eq. (4), the returned λ_j is an average of the posterior probabilities w_{ij} . In the two component models, if we modeled w_{ij} using logistic regression at the convergence of the EM, this means that λ_j becomes a function of the predictors as well as the components’ means. The final w_{ij} obtained while fitting the model described in Table 4 are used to build a logistic regression. This logistic regression models the probability of predictor vector being drawn from component number two. Then using simple algebra manipulation, we got the coefficient of the logistic model for λ_2 which are shown in Table 5. Now, the new model is exactly the model in Table 4 but with variable λ_2 and λ_1 .

Table 5. The estimated coefficient for the logistic model for λ_2 .

Predictors	Coefficient
<i>intercept</i>	-12.1702
$x_{(to-10)}^{ins}$	-0.0046
$x_{(to-5)}^{ins}$	0.0354
$x_{(to)}^{his}$	0.0890
$x_{(to+15)}^{his}$	0.0476
$x_{(to+20)}^{his}$	-0.1015
$x_{(to+25)}^{his}$	0.0718

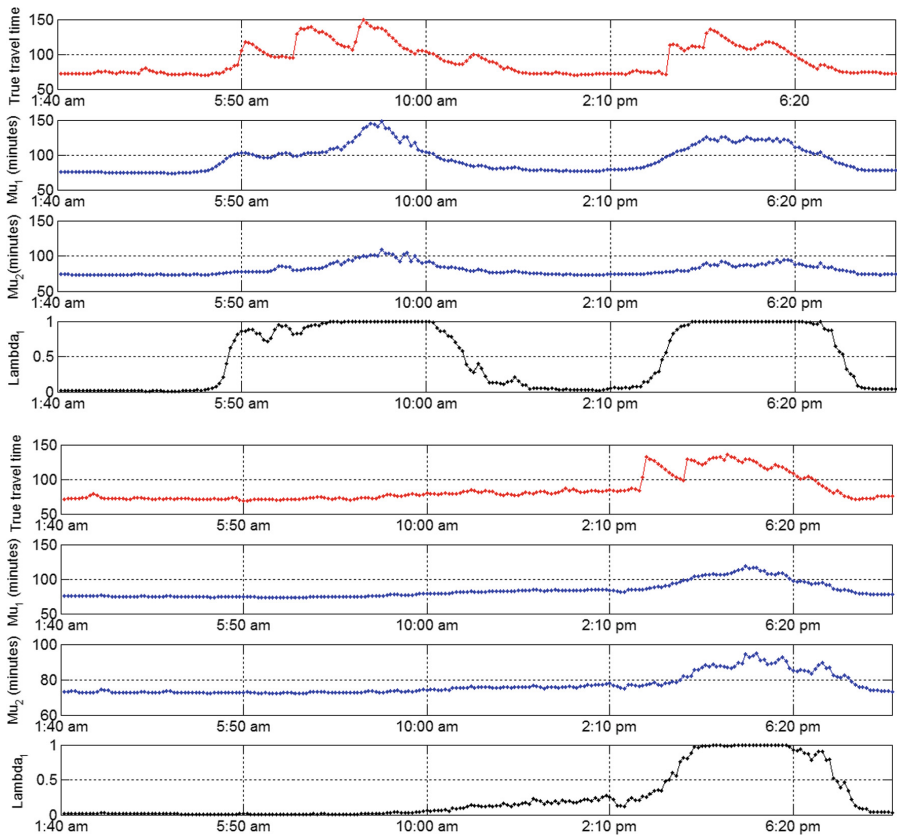


Fig. 4. The ground truth travel time (red curve), the mean of each component of the proposed model (blue curves), and the λ_1 which is the congestion probability for two different days. The upper panel is weekday and the bottom panel is a weekend day (Color figure online).

We tested the proposed model by visually inspect the ground truth travel time for each day and the mean of each component as well as the λ_1 , which is the probability of congestion in the fitted model. We visually check if the value of λ_1 is large at the time when the ground truth becomes large. As shown in Fig. 4 for weekday (top panel), there are two peaks at morning and evening and at the same time the values of λ_1 approach one which means the probability of congestion is high. The bottom panel shows a weekend where this no morning congestion but there is an evening congestion and λ_1 has only high values at the evening peak.

In order to better test the proposed model, we calculate the mean, 90% percentile, and probabilities of congestion and free flow for each predictor vector in each day of May 2013. Then based on the curves in Fig. 2, we divided the day into four time interval and calculated the mean of the above quantities within each time interval given for each day of the week. The result shown in Table 6 is consistent with the travel time pattern that we observe in Fig. 2 where at the congestion time of the day the probability of the congestion component becomes higher. Also, the model shows that the probability of the morning congestion during weekends is lower than its values at weekdays.

Table 6. Testing the model for travel time reliability using May 2013.

		1:40 am – 4:55 am		5:00 am – 10:00 am		10:05 am – 3:00 pm		3:05 am – 7:00 pm	
		Free-flow	Congested	Free-flow	Congested	Free-flow	Congested	Free-flow	Congested
Sunday	Mean	73.05	75.17	73.15	74.73	75.82	80.76	76.24	82.90
	90% percentile	75.16	85.54	75.27	85.04	78.05	91.94	78.46	94.44
	Probability	0.9853	0.0147	0.9864	0.0136	0.8769	0.1231	0.7993	0.2007
Monday	Mean	73.00	74.62	82.07	95.92	75.71	77.99	78.85	91.50
	90% percentile	75.10	84.92	84.76	109.61	78.00	88.81	81.23	104.69
	Probability	0.9876	0.0124	0.3014	0.6986	0.9014	0.0986	0.4528	0.5472
Tuesday	Mean	73.05	74.83	86.19	107.37	76.15	79.69	81.13	98.82
	90% percentile	75.16	85.16	89.02	123.02	78.46	90.81	83.60	113.27
	Probability	0.9859	0.0141	0.1140	0.8860	0.8604	0.1396	0.2819	0.7181
Wednesday	Mean	73.06	74.58	83.52	102.16	76.85	80.26	84.26	105.49
	90% percentile	75.17	84.87	86.16	117.00	79.13	91.42	86.91	120.88
	Probability	0.9874	0.0126	0.1823	0.8177	0.8525	0.1475	0.1676	0.8324
Thursday	Mean	73.02	74.58	84.54	106.41	77.61	82.37	85.91	112.97
	90% percentile	75.13	84.87	87.24	121.97	79.95	93.93	88.59	129.16
	Probability	0.9874	0.0126	0.1352	0.8648	0.7630	0.2370	0.0709	0.9291
Friday	Mean	72.88	74.37	76.93	85.48	75.78	80.43	85.92	111.09
	90% percentile	74.98	84.64	79.24	97.69	78.02	91.62	88.57	127.15
	Probability	0.9888	0.0112	0.6883	0.3117	0.8689	0.1311	0.0733	0.9267
Saturday	Mean	73.00	74.78	73.03	75.00	75.89	82.75	80.59	97.49
	90% percentile	75.11	85.09	75.14	85.36	78.11	94.19	83.04	111.24
	Probability	0.9870	0.0130	0.9837	0.0163	0.8343	0.1657	0.2593	0.7407

7 Conclusions

In this paper, we proposed a travel time model based on mixture of linear regressions. We compared two models using two different predictor sets. The first model assumes the distribution of each component in the mixture follows the normal distribution.

The second model uses the log-normal distribution instead of the normal distribution. The experimental results show that the model that uses the log-normal distribution and the historical and instantaneous travel time predictors is better than the other model. The proposed model can capture the stochastic nature of the travel time. The two-component model assigns one component to the uncongested regime and the other component to the congested regime. The means of the components are a function of various input predictors. The proposed model can be used to provide travel time reliability information at any time-of-the-day for any day-of-the-week if the predictor vector is available. The experimental results show promising performance of the proposed algorithm.

The current model does not consider the weather condition, assumes no incidents, or work zones; however this model has the ability to easily integrate these factors if the historical data includes these variables. Our future work will focus on extending, the proposed model to include these factors and study their effect on the travel time distribution.

References

1. Rice, J., van Zwet, E.: A simple and effective method for predicting travel times on freeways. *IEEE Trans. Intell. Transp. Syst.* **5**(3), 200–207 (2004)
2. Zhang, X., Rice, J.A.: Short-term travel time prediction. *Transp. Res. Part C* **11**(3–4), 187–210 (2003)
3. Kwon, J., Coifman, B., Bickel, P.: Day-to-day travel-time trends and travel-time prediction from loop-detector data. *Transp. Res. Record: J. Transp. Res. Board* **1717**(1), 120–129 (2000)
4. Chakroborty, P., Kikuchi, S.: Using bus travel time data to estimate travel times on urban corridors. *Transp. Res. Record: J. Transp. Res. Board* **1870**(1), 18–25 (2004)
5. Guin, A., Laval, J., Chilukuri, B.R.: *Freeway Travel-time Estimation and Forecasting* (2013)
6. Guo, F., Li, Q., Rakha, H.: Multistate travel time reliability models with skewed component distributions. *Transp. Res. Record: J. Transp. Res. Board.* **2315**(1), 47–53 (2012)
7. Oda, T.: An algorithm for prediction of travel time using vehicle sensor data. In: *Third International Conference on Road Traffic Control* (1990)
8. Iwasaki, M., Shirao, K.: A short term prediction of traffic fluctuations using pseudo-traffic patterns. In: *The Third World Congress on Intelligent Transport Systems, Orlando, Florida* (1996)
9. D'Angelo, M., Al-Deek, H., Wang, M.: Travel-time prediction for freeway corridors. *Transp. Res. Rec.: J. Transp. Res. Board* **1676**(1), 184–191 (1999)
10. Al-Deek, H.M., D'Angelo, M.P., Wang, M.C.: Travel time prediction with non-linear time series. In: *Fifth International Conference on Applications of Advanced Technologies in Transportation Engineering, Newport Beach, California* (1998)
11. Williams, B., Hoel, L.: Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. *J. Transp. Eng.* **129**(6), 664–672 (2003)
12. Cherrett, T.J., Bell, H.A., McDonald, M.A.: The use of SCOOT type single loop detectors to measure speed, journey time and queue status on non SCOOT controlled links, In: *Proceedings of the Eighth International Conference on Road Traffic Monitoring and Control* (1996)

13. Rilett, L., Park, D.: Direct forecasting of freeway corridor travel times using spectral basis neural networks. *Transp. Res. Rec. J. Transp. Res. Board.* **1752**(1), 140–147 (2001)
14. Matsui, H.F.M.: Travel time prediction for freeway traffic information by neural network driven fuzzy reasoning. In: Himanen, V., Nijkamp, P., Reggiani, A. (eds.) *Neural Networks in Transport Applications*. CRC Prss, Boca Raton (1998)
15. You, J., Kim, T.J.: Development and evaluation of a hybrid travel time forecasting model. *Transp. Res. Part C* **8**(1–6), 231–256 (2000)
16. Guiyan, J., Ruoqi, Z.: Travel time prediction for urban arterial road. In: *Proceedings of Intelligent Transportation Systems*. IEEE (2003)
17. Guiyan, J., Ruoqi, Z.: Travel-time prediction for urban arterial road: a case on China. In: *Proceedings of the IEEE International Vehicle Electronics Conference, IVEC 2001* (2001)
18. Wei, C.-H., Lin, S.-C., Lee, Y.: Empirical validation of freeway bus travel time forecasting. *Transp. Planning. J.* **32**, 651–679 (2003)
19. Kisgyörgy, L., Rilett, L.R.: Travel time prediction by advanced neural network. *Periodica Polytech., Chem. Eng.* **46**, 15–32 (2002)
20. Kisgyörgy, L., Rilett, L.R.: Travel time prediction by advanced neural network. *Civ. Eng.* **46** (1), 15–32 (2002)
21. De Veaux, R.D.: Mixtures of linear regressions. *Comput. Stat. Data Anal.* **8**(3), 227–245 (1989)
22. Faria, S., Soromenho, G.: Fitting mixtures of linear regressions. *J. Stat. Comput. Simul.* **80** (2), 201–225 (2009)
23. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
24. INRIX 2012. <http://www.inrix.com/trafficinformation.asp>
25. Elhenawy, M., Hassan, A.A., Rakha, H.A.: Travel time modeling using spatiotemporal speed variation and a mixture of linear regressions. In: *4th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS)*, Funchal, Madeira-Portugal (2018)