Travel Time Estimation Using Correlation Analysis of Single Loop Detector Data

Submission Date: Nov 10, 2005

Word Count: 7338

Authors:

Huairui Guo
Research Scientist
Reliasoft Corporation,
115 South Sherwood Village Drive
Tucson, AZ, 85710
Tel: 520-886-0366
Fax: 520-886-0399
Email: Harry.Guo@Reliasoft.com

Jionghua (Judy) Jin
Associate Professor
Industrial and Operations Engineering Department
University of Michigan,
Ann Arbor, MI 48109-2117
Tel: (734) 763-0519
Fax: (734) 764-3451
Email: jhjin@umich.edu

Travel Time Estimation Using Correlation Analysis of Single Loop Detector Data

Abstract: In this paper, the average travel time over a link is considered as a random variable following an identical probability distribution as the arrival process. A new estimation method of the average travel time was proposed by using the cross-correlation analysis of traffic flow measurement data. By using the proposed estimation method, it only requires the traffic flow information, which is available through the measurements of single-loop detectors at the upstream and downstream of one link. Different from the existing maximal cross-correlation analysis method, this paper considers the average travel time as a random variable with its mean value estimated from all significant cross-correlation coefficients rather than only using the maximum cross-correlation coefficient. Therefore, the inherent variability of the average travel time among different vehicles can be considered. Moreover, different from the existing optimization method, the statistical t-test of the significant cross correlation coefficients is used to automatically and adaptively determine the fitting range of the probability density function of the average travel time. Thus, it avoids using the approximated car length factor and has no need to pre-determine the range of the average travel time as required by the optimization method. In this paper, the details of the average travel time estimation procedures are presented and the effectiveness of the proposed method is demonstrated through both simulation study and a case study of real traffic data.

Keyword: travel time estimation, correlation analysis, single loop detector, t-test

1. INTRODUCTION

Accurate estimation of link travel time has essential importance in the development of intelligent transportation system. The travel time estimation provides not only valuable information for traveler routing or transportation scheduling, but also the potential capability for incident detection. Currently, single-loop detectors have been widely used in arterial roads and freeways for traffic condition monitoring and control. Generally, such detector data can be used to provide a direct measurement of traffic flow but cannot provide vehicle speed and travel time information.

Various models have been developed for travel time estimation. Among them, the most simple and fundamental model is to use the basic relationship among speed, volume and occupancy to estimate the travel time (*1, 2, 3*). The basic idea of this model is briefly reviewed as follows.

For a given link, the relationship of average speed, flow, and occupancy is expressed as:

$$v = \frac{q}{o \cdot g} \tag{1}$$

where v is the speed, q is the flow, o is the occupancy, and g is the car length factor that is a function of the lengths of vehicles and detectors. The average travel time is estimated by:

$$T = \frac{s}{v} \tag{2}$$

where s is the length of a given link.

Although the model of Equations (1) and (2) provides a simple way for average travel time estimation, its estimation accuracy is very sensitive to the car length factor g, which will be further elaborated later in this paper. In addition, the travel time variability among different vehicles is not considered in this model.

Recently, some other techniques were developed for travel time estimation. As an example, Coifman used instant velocity measurement by a dual loop detector (*4*). This method is effective if dual loop detectors have been equipped in a link. However, it is not suitable for the links that are only equipped with single-loop detectors because single-loop detector cannot measure the speed directly. Another technique is to use cumulative flow plot to calculate total travel delay and to estimate average travel time (*5, 6*). This technique requires the priori knowledge about the number of vehicles initially existing in the link, which is usually not available in practice. In addition, there is no systematic way to choose the start time to plot the departure curves. Thus, the implementation of this technique is limited. Nam and Drew (*7*) gave a detail analysis of using cumulative flow plot to estimate the travel time. They also presented the new development of the flow-density-speed relationship of their models. However, it still has the same problem as references (*5, 6*).

Some different methods were developed for travel time estimation by applying advanced signal processing and intelligent classification techniques such as state-space modeling, neural network classification, and car identification and signature matching. Park et al. used neural networks to forecast real-time travel time for 5-mins ahead (*8*). Hoogendoorn et al. proposed a method that describes state space dynamics of traffic with recurrent neural networks (*9*). Both

methods require appropriate selection of training data sets to conduct extensive training of the neural network models. Coifman and Cassidy developed an algorithm for travel time estimation through vehicle re-identifications, in which the vehicle length is used as a signature to match a vehicle in the downstream to the vehicle in the upstream (*5*). Thus, the vehicle length has to be measured or estimated accurately to ensure the vehicles to be re-identified correctly.

In order to consider the stochastic nature of traffic, some stochastic models of traffic flow were developed in the last decades. Dailey used cross-correlation analysis for travel time estimation, which assumes that the average travel time is a deterministic variable (*10*). This method only used the maximal cross-correlation coefficient to estimate the average travel time. Thus, it is effective when different vehicles have a closer travel speed. Recently, Dailey modifies the model of Equation (1) by considering the measurement errors of the car length and occupancy (*11*). The modified model considered the occupancy, speed, and vehicle length as random variables following certain probability distributions. Then, a Kalman filter is used to estimate the actual speed of each vehicle. This method requires a prior understanding of the probability distribution function of various random variables involved in the model. A significant estimation error, even instability or non-convergence, may occur if the mean or variance of the random variables were not properly estimated.

In addition, by considering the random characteristics of average travel times, Petty et al developed a probabilistic regression model with the assumption that the average travel time is a random variable with the same probability distribution function over a given link for a given estimation window of travel time (*12*). Although this method is robust to congestion conditions, the effectiveness of this method is highly dependent on the prior knowledge of the fit range of the probability density function of the average travel time. In their paper, the center of the fit range is adaptively estimated based on Equation (1) with a fixed value of g. The change of the car length factor due to the instant change of occupancy is not considered, which could lead to an estimation error on the fit center of the probability density function. Moreover, the selection of a fit range will also affect the estimation performance. The influence of these two factors on the average travel time estimation will be further illustrated later in the paper.

In summary, the existing travel time estimation methods limited by using the traffic flow data from single loop detectors can be classified as three categories: one is rooted from the model of Equation (1) for travel speed estimation, which is sensitive to the car length factor estimation error. The second method is to consider the stochastic nature of the traffic flow and use the maximum cross correlation coefficient between the flow data at the upstream and down stream to estimate the average travel time, which ignores the travel time variability among different vehicles. The third method is to develop a probabilistic regression model to estimate the probability density function of the average travel time, which requires the prior determination of the fit range of the estimated probability density function of the average travel time. Therefore, this paper aims to improve the existing travel time estimation methods based on the flow measurement data from single loop detectors at the upstream and the downstream.

In this paper, a new method is developed for travel time estimation by integrating the cross correlation analysis technique with the probabilistic model of the random average travel time. In the paper, similar to the Petty's method, a probabilistic regression model is used to consider the randomness of the average travel time. The advantage of the proposed analysis method is to improve its estimation method by automatically determining the fit range of the probability density function of the average travel time adaptively. Moreover, this proposed

estimation method does not require the estimation of car length factor that may lead to significant estimation errors. The only information needed in the proposed method is the flow measurement at downstream and upstream. In the following sections, Section 2 will briefly review the probabilistic regression model and discuss the impact of the car length factor estimation errors on the existing travel time estimation method. In Section 3, a new cross correlation analysis method is proposed to consider the travel time variability among different vehicles. A statistical hypothesis testing is used to automatically determine the fit range of the probability density function of the average travel time. In Section 4, both simulation study using the VISSIM software and a case study of real traffic data will be conducted to demonstrate the analysis procedures and the effectiveness of the proposed method. Finally, the paper is summarized and concluded in Section 5.

2. REVIEW OF TRAVEL TIME ESTIMATION METHODS USING FLOW MEASUREMENTS

2.1 Review of Probabilistic Regression Model of Flow Measurements

The travel flow model used in this paper is based on the probabilistic regression model proposed by Petty et al (*12*). It assumes that the arrivals measured at the upstream point within a given time interval (called estimation window) have the same probability density function of the average travel time over the link. If at sampling time t, the counts of arrivals during the sampling interval Δ is $x(t)$ and $y(t)$ corresponding to the upstream and downstream respectively, the relationship of these upstream and downstream can be modeled as:

$$y(t) = \sum_{i=a_1}^{a_2} [x(t-i)\, f_i] \tag{3}$$

where f_i is the probability density function of the average travel time over the link. a_1 and a_2 correspond to the shortest and longest average travel time over all upstream vehicles respectively, and $[a_1, a_2]$ is called fit range of the probability density function. In (*12*), the estimation of f_i is obtained to minimize the sum of squares of the regression residual errors by solving the following constrained optimization problem.

$$\min_{f_i} \{ \sum_{t=b_1}^{b_2} (y(t) - \sum_{i=a_1}^{a_2} x(t-i)\, f_i)^2 \} \tag{4}$$

$$\text{Subject to: } \{ \mathbf{f}: f_i \geq 0, \sum_{i=a_1}^{a_2} f_i = 1 \}$$

where b_1 and b_2 are the start time index and end time index of the estimation window $[b_1, b_2]$. Thus, the mean estimate of the average travel time is:

$$\hat{T} = \Delta \sum_{i=a_1}^{a_2} f_i \cdot i \tag{5}$$

From Equation (4), it can be seen that if the predefined fit range of $[a_1, a_2]$ is too wide,

there will be too many parameters f_i to be estimated, which may lead to an unstable solution in the optimization solver. Meantime, if the fit range is too narrow, the resolution and accuracy of the estimated probability density function may be too poor to get an appropriate probability distribution function. Therefore, it is very critical to define an appropriate fit range of [a_1, a_2]. As the paper of Petty et al. (*12*), the fit range is fixed as 20 seconds independent of the occupancy change, and the center of the mean travel time is adaptively estimated based on Equation (2) with a fixed car length factor of g in Equation (1). The following analysis will illustrate how this fixed car length assumption affects the fit center of [a_1, a_2] used in the travel time estimation of Equation (4).

2.2 Impact Analysis of Car Length Factor Estimation Error

Hall and Persaud made a thorough discussion about the relationship among speed, occupancy and car length factor g (*2*). The results showed that there is a strong correlation between occupancy and the average car length. With the increase of occupancy, the average car length also increases. So, the car length should be adaptively estimated for different occupancies. Now, we will further illustrate how the estimation error of using a fixed car length factor affects the accuracy of travel time estimation.

From Equations (1) and (2), the fit center of [a_1, a_2] used in the Petty's estimation method is obtained as:

$$\hat{T}_c = \frac{s \cdot o}{q \cdot l} \tag{6}$$

Here, $l = 1/g$ that represents the average effective car length, or the sum of the average car length and the width of the loop detector. When dual loop detectors are used, the traffic flow (q) and the traffic occupancy (o) are directly obtained and adaptively used to estimate the travel time ($\hat{T}_c$), which is used as the center of the fit range of [a_1, a_2] in every estimation window of [b_1, b_2]. In this case, a fixed car length is used in Equation (6). Based on Equation (6), the resulted deviation error Δl due to incorrectly using the fixed average car length can cause the estimation error ΔT_c as:

$$\Delta T_c = -\frac{s \cdot o}{l^2 \cdot q} \Delta l \tag{7}$$

The impact of Δl on the ΔT_c is further investigated through a simulation study by using the VISSIM simulation software. The conditions used in the simulation study are set as follows: link distance $s = 300m$, average traffic flow q = 1000 *vehicle/hour* (0.28 *vehicle/second*). The estimation window is one minute. For each estimation window, the occupancy and the corresponding average car length are calculated from simulation data. The relationship between the occupancy and average car length is plotted in Figure 1. From Figure 1, it is found that the overall average of the car length estimation is 6.52*m*, and the corresponding occupancy is 0.12**.** If we divide Equation (7) by Equation (6), it has:

$$\frac{\Delta T}{\hat{T}_c} = \frac{-\Delta l}{l} \tag{8}$$

FIGURE 1 Correlation between average car length and occupancy.

From Equation (8), it can be seen that the estimation error of the average travel time is proportional to the estimation error of the average car length. If the overall average car length of 6.52*m* is used for the value of l in Equation (6), the range of $\Delta l / l$ will be [-0.29 0.36] in the above simulation. The effect of the estimation error of the average car length on the estimated average travel time $\hat{T}_c$ is obvious. So, if this inaccurate value of $\hat{T}_c$ is further used as the center of the fit range $[a_1, a_2]$, a biased estimate of the average travel time could be generated because the predefined fit width $\Delta=a_2-a_1$ may not be sufficiently cover all possible travel time values under different occupancy conditions. The influence of using a fixed car length on the travel time estimation will be further illustrated through a simulation analysis in Section 4.3.

In the following sections, a new method is proposed to automatically determine the fit range of $[a_1, a_2]$ based on the statistical correlation analysis and t-test.

3. PROPOSED CORRELATION ANALYSIS METHOD FOR TRAVEL TIME ESTIMATION

3.1 Correlation Analysis of Upstream Flow and Downstream Flow

Two data series of $x(t)$ and $y(t)$ are used to denote the travel flow at upstream and downstream respectively. The cross covariance of $x(t)$ and $y(t)$ and the autocovariance of $x(t)$ with lag k are defined as:

$$\gamma_{yx}(k) = E[(y(t) - u_y)(x(t-k) - u_x)] \tag{9}$$

$$\gamma_x(k) = E[(x(t) - u_x)(x(t-k) - u_x)] \tag{10}$$

Thus, the cross correlation coefficient and autocorrelation coefficients are:

$$\rho_{yx}(k) = \frac{\gamma_{yx}(k)}{\sigma_x \sigma_y} \tag{11}$$

$$\rho_x(k) = \frac{\gamma_x(k)}{\sigma_x^2} \tag{12}$$

where σ_x and σ_y are the standard deviation of $x(t)$ and $y(t)$ respectively. It is assumed that $x(t)$ and $y(t)$ are stationary series within the estimation window of $[b_1, b_2]$. From Equation (3), it has:

$$u_y = \sum_{i=a_1}^{a_2} f_i u_x \tag{13}$$

where $u_y = E(y_t)$ and $u_x = E(x_t)$. Based on Equations (3), (9)-(13), it has:

$$\rho_{yx}(k) = \frac{1}{\sigma_x \sigma_y} \cdot E[(\sum_{i=a_1}^{a_2} f_i x(t-i) - \sum_{i=a_1}^{a_2} f_i u_x)(x(t-k) - u_x)]$$
$$= \frac{1}{\sigma_x \sigma_y} \cdot \sum_{i=a_1}^{a_2} f_i \gamma_x(i-k) \quad (14)$$
$$= \frac{\sigma_x}{\sigma_y} \cdot \sum_{i=a_1}^{a_2} f_i \rho_x(i-k)$$

It can be seen that the unknown parameters f_i ($i \in [a_1, a_2]$) can be estimated through the correlation analysis of the upstream and downstream flow.

It should be clarified that although the proposed approach uses the correlation analysis to estimate the travel time, there is a major difference from the Dailey's maximal cross-correlation method (*10*). As we reviewed in Section 1, Dailey's maximal cross-correlation method assumed the average travel time is one fix value and used only the maximum cross correlation coefficient for the average travel time estimation. Thus, it does not consider the variability of average travel time due to different travelers. In the proposed method, we assume the average travel time is a random variable over different travelers and use Equation (14) to estimate a set of f_i as the discretized probability density function by the sampling interval Δ, in which multiple significant correlation coefficients of $\rho_{yx}(k)$ will be used as follows.

Under the free flow condition, it is assumed that $x(t)$ at upper stream are independent samples, thus

$$\rho_x(i-k) = \begin{cases} 1 & i = k \\ 0 & others \end{cases} \quad (15)$$

Based on Eqs. (14) and (15), it can be obtained as:

$$f_k = \frac{\sigma_y}{\sigma_x} \rho_{yx}(k) \quad (16)$$

From Equation (16), it can be seen that f_k can be solely determined by the cross correlation coefficients of $\rho_{yx}(k)$. The task of estimating f_k is transferred to find the nonzero cross correlation coefficients, which is simpler and straightforward. The proposed method does not need to predetermine the car length factor and the fit ranges of $[a_1, a_2]$ for f_i. Moreover, the proposed method using Equation (16) can avoid the unstable solution problem that occurs in the optimization approach.

3.2 Hypothesis Testing of Cross-Correlation Coefficients

The aforementioned analysis indicates that the discretized average travel time distribution can be estimated directly from the nonzero cross-correlation coefficients. In practice, the true value of $\rho_{yx}(k)$ is unknown. Thus, the sample cross correlation function $\hat{\rho}_{yx}(k)$ is used, which is calculated from the flow data within the estimation window [b_1, b_2].

$$\hat{\rho}_{yx}(k) = \frac{\sum_{t=b_1+k}^{b_2}(x(t-k)-\bar{x})(y(t)-\bar{y})}{\sqrt{\sum_{t=b_1}^{b_2-k}[(x(t)-\bar{x})^2][\sum_{t=b_1+k}^{b_2}(y(t)-\bar{y})^2]}} \tag{17}$$

where $\bar{x} = [\sum_{t=b_1}^{b_2} x(t)]/(b_2\text{-}b_1+1)$, $\bar{y} = [\sum_{t=b_1}^{b_2} y(t)]/(b_2\text{-}b_1+1)$.

In order to determine the nonzero cross correlation coefficients, the following statistical hypothesis test is used, that is,

$$\begin{aligned} &H_0: \ \rho_{yx}(k) = 0; \\ &H_1: \ \rho_{yx}(k) \neq 0; \end{aligned} \tag{18}$$

When the estimation window is appropriately selected to include enough number of data, $\hat{\rho}_{yx}(k)$ under H_0 approximately follows a normal distribution with mean zero and the standard deviation $\sigma_{\hat{\rho}}$ is (*13*):

$$\sigma_{\hat{\rho}} = \frac{1}{\sqrt{n}} \tag{19}$$

where $n = b_2\text{-}b_1 - k+1$ is the number of data points used in the estimation of $\hat{\rho}_{yx}(k)$. Thus, a t-test can be used to check if $\rho_{xy}(k)$ is equal to zero under a given α value. The statistic $t_{\rho,k}$ is defined as:

$$t_{\rho,k} = \frac{\hat{\rho}_{yx}(k)}{\sigma_{\hat{\rho}}} \tag{20}$$

Under the given α error, if the condition of $t_{\rho,k} > t_{1-\alpha/2,n}$ holds, we reject H_0. Thus, the cross correlation at lag k is significant, and f_k calculated from Equation (16) is the estimated probability corresponding to the average travel time $k \cdot \Delta$.

It should be clarified that under the severe congestion condition at the period of the estimation window, there will be no significant cross-correlation coefficients found. In fact, in the congestion condition, it is more important to conduct traffic condition monitoring rather than travel time estimation. The lack of the significant cross correlation coefficients will be used an effective indicator for detecting the congestion condition. An example of how to automatically determine the significant cross correlation coefficients will be shown in Section 4.

The proposed model is based on the cross-correlation values of upstream and downstream

volume which requires the real time flow information of each sampling interval. Traditional flow-density models do not have this information and thus cannot be used to analytically validate the proposed model. Therefore, we used VISSIM to conduct a series of simulations and studied how the traffic volume will affect the cross-correlation values. A 300 meter long link with capacity 1900 vehicle/ hr was used. The estimation window is 10 minutes. According to the above analysis, $t_{0.95,600} = 1.96$. Therefore the threshold value of rejecting H_0 is 0.08. The setting of the traffic flow starts from 200 vehicle/hr and ends at 1900 vehicle/hr with increasing step 100 vehicle/hr. For each setting, the simulation lasted 1 hour and the average of all the maximal cross-correlation values of each estimation window is calculated and plot in Figure 2.

Under the free flow condition, the increasing of the traffic volume can lead to the decrease of the maximal cross-correlation coefficients. However, for all the volumes that are less than the road capacity, the cross-correlation values are larger than 0.08, which means there is a strong cross-correlation between the upstream and downstream traffic flows. Therefore, the proposed method can be effectively used for travel time estimation under the free flow condition.

FIGURE 2 Cross-Correlation between x and y under different flow conditions.

3.3 Mean Estimate of Average Travel Time

The estimation of f_i using Equation (16) only considers those non-zero probability values of f_i which are corresponding to a certain percentage of vehicles having the same average travel time over the link. The summation of all estimated f_i will represent the percentage of the upstream vehicles within the estimation window of $[b_1, b_2]$ used for the average travel time estimation, that is, $\sum_{i \in c_1}^{c_m} f_i = P$. Therefore, among the upstream vehicles used for the travel time estimation, $100\ f_i / P\%$ of vehicles have the travel time of $c_i \cdot \Delta$, i=1, …, m. So, the mean estimate of the average travel time based on these $100P\%$ upstream vehicles is:

$$\hat{\mu}_T = \Delta \cdot \sum_{i=c_1}^{c_m} [(f_i / P) \cdot i] \tag{21}$$

It can be seen that the mean estimate using Equation (21) ignores the effect of the small probability of the vehicles having either extremely long travel time (or never arriving at the downstream detector) or extremely short travel time, which provides more realistic value of the mean estimate of the average travel time. By substituting the sample cross correlation function of Equation (17) into Equation (21), we can obtain the mean estimate of the average travel time as:

$$\hat{\mu}_T = \Delta \cdot [\sum_{i=c_1}^{c_m} \hat{\rho}_{yx}(i) \cdot i)] / \sum_{i=c_1}^{c_m} \hat{\rho}_{yx}(i) \tag{22}$$

4. SIMULATION STUDY AND CASE STUDY

4.1 Description of Simulation Conditions

Four simulation studies were conducted, which will be discussed in the following four subsections. In those simulation studies, VISSIM simulation software is used to generate the travel flow data and obtain the reference of the true travel time values. The total simulation time is five hours but we only use the data generated in the last 4 hours for the analysis. The common conditions used for those four simulation studies are listed as follows:

- Traffic flow: q=1000vehicles/hour
- Speed: v=45km-60km/hour (12.5-16.67 meters/second)
- Link distance s=300 meters
- Data sampling interval of flow measurement Δ=1 sec
- The width of the estimation window b_2-b_1=10 mins

4.2 Cross Correlation Analysis and t-test

This simulation study is used to illustrate the procedures of the cross correlation analysis and the t-test discussed in Section 3.2. As an example, the upstream and the downstream data are selected from an arbitrary estimation window. The sample cross correlation function $\hat{\rho}_{yx}(k)$ ($0<k\leq 35$) are calculated based on Equation (17) as shown in Figure 3. Since $n = b_2 - b_1 - k + 1$ is very large, the value of $t_{1-\alpha/2, b_2-b_1-k+1}$ will not depend on the lag k. If α error is selected as 5%, we have $t_{1-\alpha/2, b_2-b_1-k+1} = t_{0.95, 600-k+1} = 1.96$. So, based on Equation (20), the threshold η for determining the significant cross correlation coefficients is obtained as:

$$\eta = \frac{t_{0.95,600-k+1}}{\sqrt{n}} = \frac{1.96}{\sqrt{600-k+1}} \approx \frac{1.96}{\sqrt{600}} = 0.08$$

From Figure 3, it is easy to find that the cross-correlation coefficients from lag 20 to lag 25 are significant because they are larger than the threshold 0.08. So, based on Equation (22), the mean estimate of the average travel time over the link is obtained as:

$$\hat{\mu}_T = \Delta \cdot \sum_{i=20}^{25} \{[\hat{\rho}_{xy}(i)] \cdot i\} / \sum_{i=20}^{25} [\hat{\rho}_{xy}(i)] = 22.35$$

Based on above analysis, it can be seen that by using t-test, the proposed estimation method has no requirement on the pre-selection of the fit range of $[\,a_1 \;\; a_2\,]$, and no need of using the optimization solver.

FIGURE 3 Cross-correlation analyses between the upstream and downstream.

4.3 Comparison of Different Travel Time Estimation Methods

This simulation is used to compare the performance of different travel time estimation methods. As is discussed before, two parameters need to be set for using the Petty's optimization method. One parameter is the width of the fit range equal to $a_2 - a_1 = 20$sec. The other one is the fixed average car length used in Equation (6) for the estimation of the center of the fit range. For comparing the optimal performance of Petty's method (ignoring the bias in the estimation of average car length), the true average car length of 6.5 meter, which is obtained from VISSIM software, is used in the simulation. The travel time estimation results, which are obtained by using (1) the optimization method, (2) single (maximal) cross correlation coefficient, and (3) the proposed method, are all compared with the reference of the true travel time as shown in Figure 4.

For further comparison, the estimation errors of each method (i=1,2,3) are calculated by the deviation e_i of the estimated travel time $\hat{\mu}_{T_i}$ from the true travel time $\hat{\mu}_{T_0}$, that is, $e_{i=} \hat{\mu}_{T_i} - \hat{\mu}_{T_0}$. The average and the standard deviation of the estimation errors of each method are compared in Table 1. From Table 1, it can be seen that the proposed method provides a better performance than other two methods.

FIGURE 4 Comparison of different travel time estimation methods.

TABLE 1 Comparison of the Estimation Errors

4.4 Effect of Estimation Window Length on Different Travel Time Estimation Methods

This simulation is used to study how the estimation window length affects the performance of each estimation method. Five different estimation window lengths (b_2-b_1) are used in the simulation, which are b_2-b_1= 90, 150, 300, 450, 600 (sec). The other simulation parameters are kept the same as in Section 4.4. The comparison of the three travel time estimation methods are plotted in Figure 5. Figure 5 (a) and (b) are corresponding to the average of the estimation errors and standard deviation of the estimation errors respectively.

FIGURE 5 Comparison of the estimation errors under different fitting window lengths.

From Figure 5, it is clear that for all estimation methods, the standard deviations of the estimation errors are significantly reduced when the estimation window length is equal to or larger than 300 secs. However, the selection of too large estimation window will lose the dynamic behavior of the traffic condition change. So, an appropriate selection of the estimation window length is important. Figure 5 shows that under the given simulation condition, 600 sec is a relatively good estimation window length for all three estimation methods.

When comparing the maximal cross-correlation method with the proposed method, Figure 5 shows that when the estimation window length is less than 300sec, both methods have

the similar estimation performance. However, with the increase of the estimation window length, the proposed method shows a better performance than the maximal cross-correlation method. This is because with the increase of the estimation window length, more upstream arrivals having different travel time are included. The maximal cross correlation method relying only on a single maximal cross correlation coefficient cannot consider the traveler variability and also will be more sensitive to the noise influence on determining the maximum cross correlation coefficient. The proposed method using all significant cross-correlation coefficients for the travel time estimation can consider the traveler variability with different travel times. Moreover, the use of the multiple significant cross correlation coefficients can smooth the estimation error of the cross correlation coefficients. That is why with the increase of the estimation window length, the proposed estimation method is better than the maximal cross-correlation method.

4.5 Case Study

This case study is used to illustrate how the proposed method is used for travel time estimation in a real traffic condition. The flow data used in this study is collected by the single loop detectors at Speedway Blvd in Tucson city from 11:30AM to 12:30AM on 12/13/200, in which single loop detectors are used to measure the traffic flow at each direction. As an example, Figure 6 shows the detector locations in the intersection of Speedway Blvd and Tucson Blvd, where the data are collected for this case study. There are three lanes in the Speedway Blvd. The traffic data of the middle lane from Tucson Blvd to Country Club Blvd collected by detector 1 was used for travel time estimation. The link distance is 2689 feet (817 meters), and the speed limit is 35 miles (15.6meters/sec). In this case study, our travel time estimation is updated for every 10 mins. Figure 7 shows the cross correlation analysis results in each estimation window. The travel time estimation results are shown in Figure 8. A reasonable increasing trend of the estimated travel time is observed when it approaches to noon for the lunch break.

FIGURE 6 Single loop detector locations in Speedway Ave at Tucson.

FIGURE 7 Cross-correlation analyses at different time periods.

FIGURE 8 Estimated average travel time using the proposed method.

5. CONCLUSION

This paper proposed a new travel time estimation method. By using the proposed correlation analysis method, it only needs the travel flow information measured by single loop detectors. Therefore, the estimation errors caused by the measurement errors of occupancy, vehicle lengths are avoided. By using the t-test, the proposed method can automatically determine the center and width of the estimation window of the probability distribution function of average travel time. Several comparison analyses between the proposed method and two existing methods were conducted through the simulation studies. A case study also shows the effectiveness of the proposed method for the real traffic data analysis.

ACKNOWLEDGEMENT

The authors would like to thank the reviewers for their insightful comments and suggestions, which have significantly improved the paper quality and readability. The authors also gratefully acknowledge the financial support of the U.S. Department of Transportation/Bureau of Transportation Statistics (DOT/BTS) Grant: DTTS-01-G-B101-AZ and DTTS-02-G-B201-AZ; the Federal Highway Administration/Arizona Department of Transportation (FHWA/ADOT) Grant: JPA0281; and the National Science Foundation (NSF) Grant: PECASE/CAREER AWARD DMI-0133942.

REFERENCES

[1] Athol, P. Interdependence of certain operational characteristics within a moving traffic stream. *Highway Research Record*, 72, 1965, pp. 57-87.

[2] Hall, F. L. and Persaud, B. N. Evaluation of speed estimation made with single-detector data from freeway traffic management systems. *Transportation Research Record*, 1232, 1989, pp. 9-16.

[3] Jacobson, L. N., Nihan, N. L. and Bender, J. D. Detecting erroneous loop detector data in a freeway traffic management system. *Transportation Research Record* 1287, 1990, pp. 151-166.

[4] Coifman, B. Estimating travel times and vehicle trajectories on freeways using dual loop detectors. *Transportation Research 36A*, 2002, pp. 351-364.

[5] Coifman, B. and Cassidy, M. Vehicle reidentification and travel time measurement on congested freeway. *Transportation Research 36A*, 2002, pp. 899-917.

[6] Nam, D. H. and Drew, D. R. Traffic dynamics: method for estimating freeway travel times in real time from flow measurements. *Journal of Transportation Engineering*, 122, 1996, pp. 185-191.

[7] Nam, D. H. and Drew, D. R. Automatic measurement of traffic variables for intelligent transportation system applications. *Transportation Research 33B*, 1999 , pp. 437-457.

[8] Park, D., Rilett, L R. and Han, G. Spectral basis networks for real-time travel time forecasting." *Journal of Transportation Engineering*, 122, 1996, pp. 185-191.

[9] Hoogendoorn, S. P., Van Lint, J.W.C. and Van Zuylen, H. J. Freeway travel time prediction with State-Space neural network. Preprint 02-2797, *Proceedings of 81st Annual Meeting*, Transportation Research Board, Washington, D. C. Conference CD. 2001.

[10] Dailey, D. J. Travel time estimation using cross-correlation techniques. *Transportation Research* 27B, 1993, pp. 97-107.

[11] Dailey, D. J. A statistical algorithm for estimating speed from single loop volume and occupancy measurement. *Transportation Research 33B*, 1999 , pp. 313-322.

[12] Petty, K., Bickel, P., Jiang, J., Ostland, M., Rice, J., Ritov, Y., and Schoenberg, F. Accurate estimation of travel times from single-loop detectors. *Transportation Research 32A*, 1998 ,pp. 1-17.

[13] Shumway, R. H. and Stoffer, D. S. *Time Series Analysis and Its Applications.* Springer-Verlag, New York. 1999.

List of Figures and Tables

TABLE 1 Comparison of the Estimation Errors
FIGURE 1 Correlation between average car length and occupancy.
FIGURE 2 Cross-Correlation between x and y under different flow conditions.
FIGURE 3 Cross-correlation analyses between the upstream and downstream.
FIGURE 4 Comparison of different travel time estimation methods.
FIGURE 5 Comparison of the estimation errors under different fitting window lengths.
FIGURE 6 Single loop detector locations in Speedway Ave at Tucson.
FIGURE 7 Cross-correlation analyses at different time periods.
FIGURE 8 Estimated average travel time using the proposed method.

TABLE 1 Comparison of the Estimation Errors

Method / Value	(1)Optimization method	(2)Single cross-correlation method	(3)Proposed multi-cross-correlation Method
Mean of error	0.2464	0.4547	0.2427
Standard Deviation of error	1.238	1.1759	0.4814

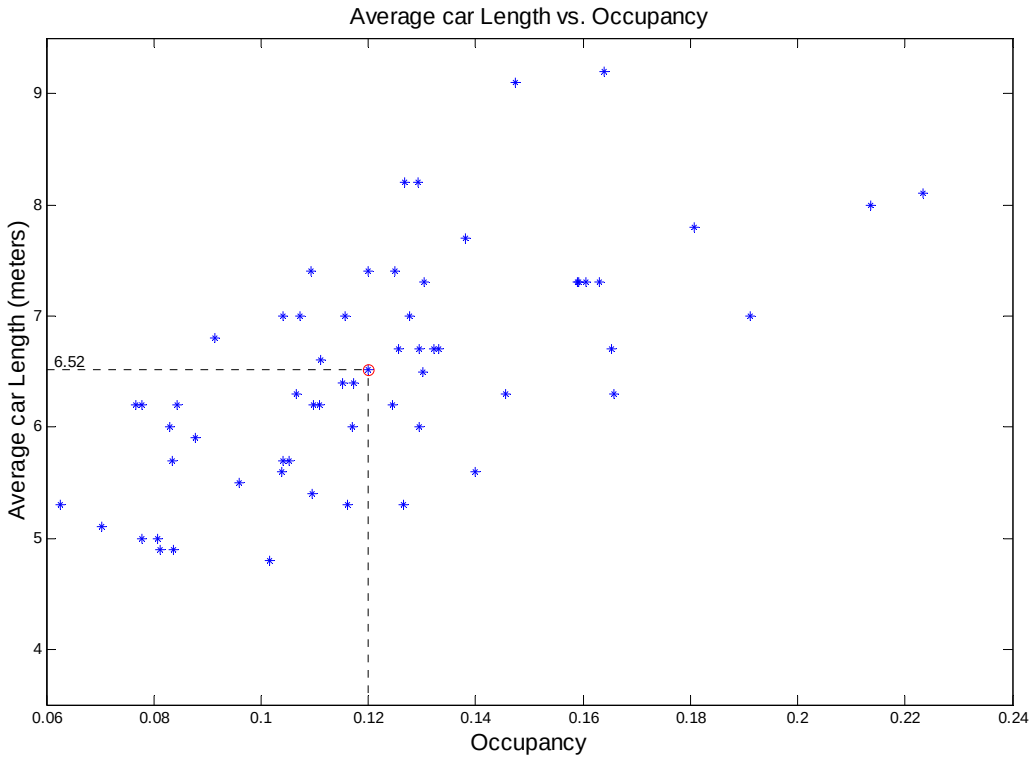


FIGURE 1 Correlation between average car length and occupancy.

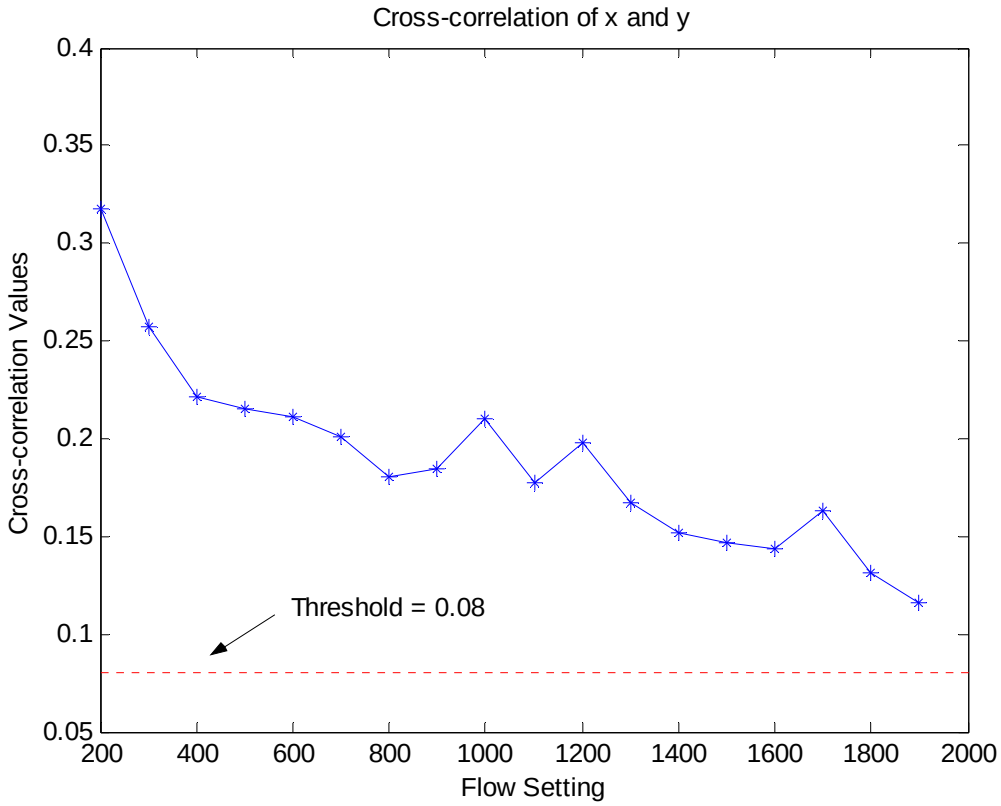


FIGURE 2 Cross-Correlation between x and y under different flow conditions.

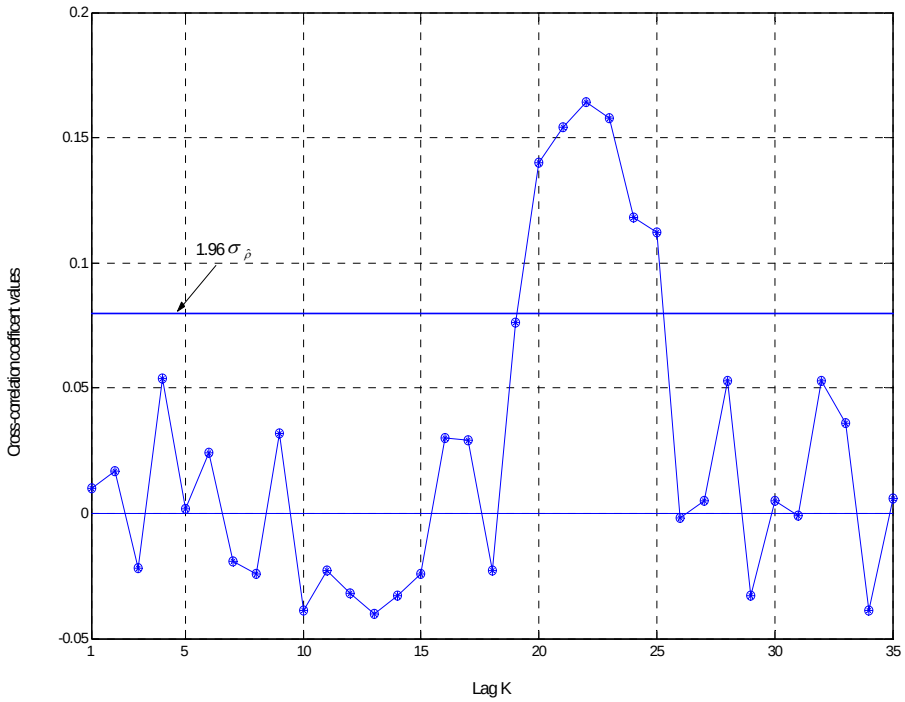


FIGURE 3 Cross-correlation analyses between the upstream and downstream.

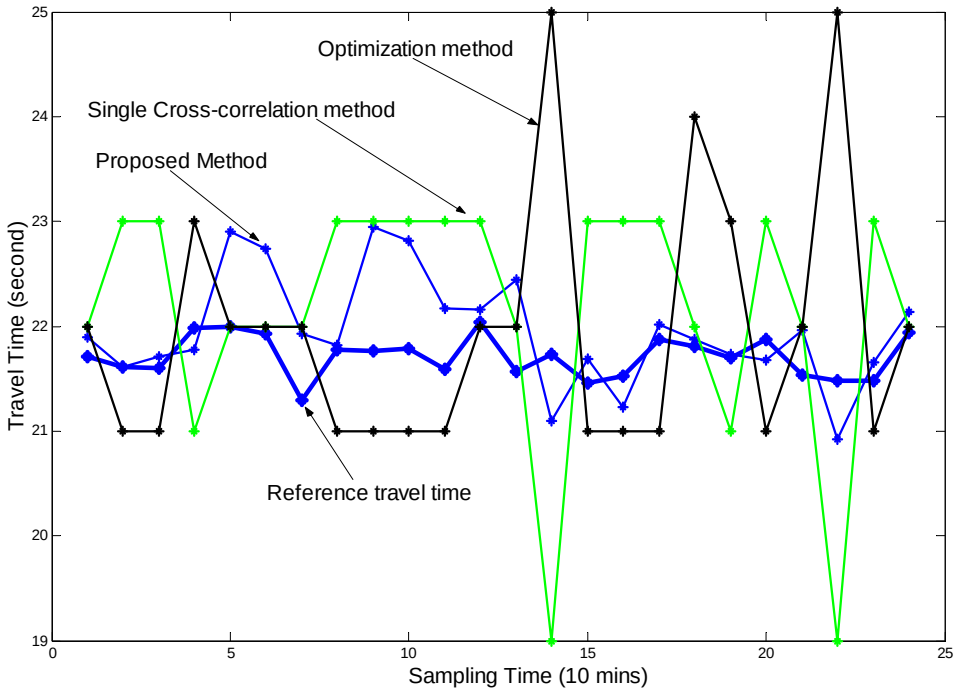


FIGURE 4 Comparison of different travel time estimation methods.

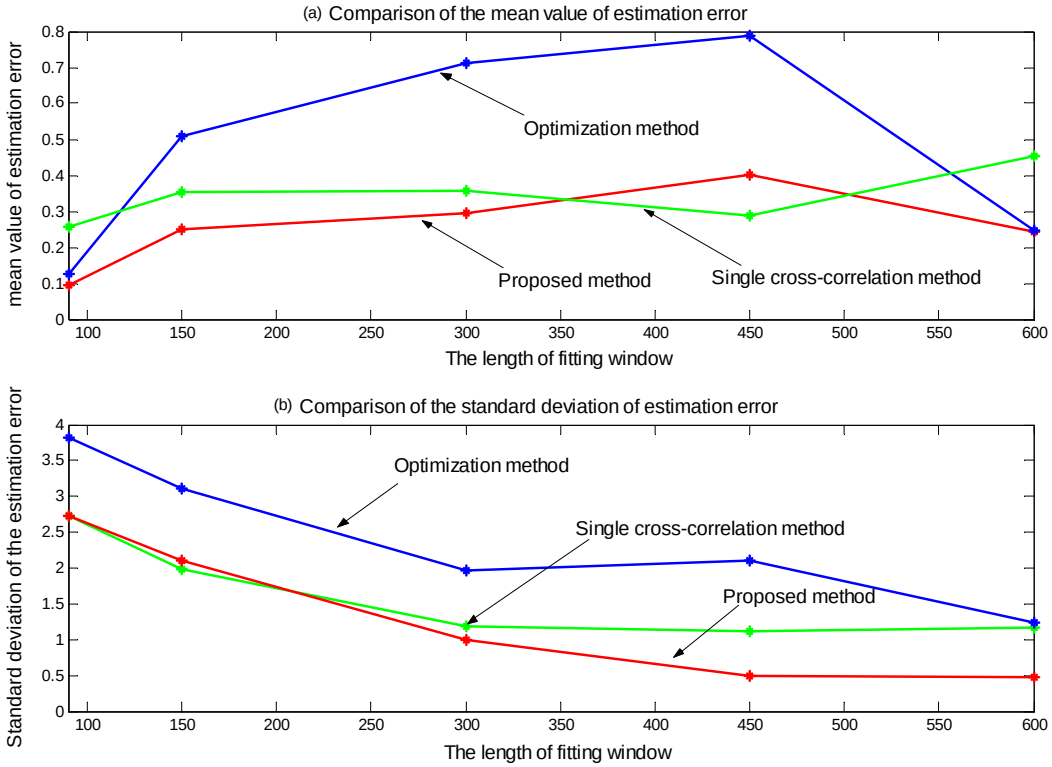


FIGURE 5 Comparison of the estimation errors under different fitting window lengths.

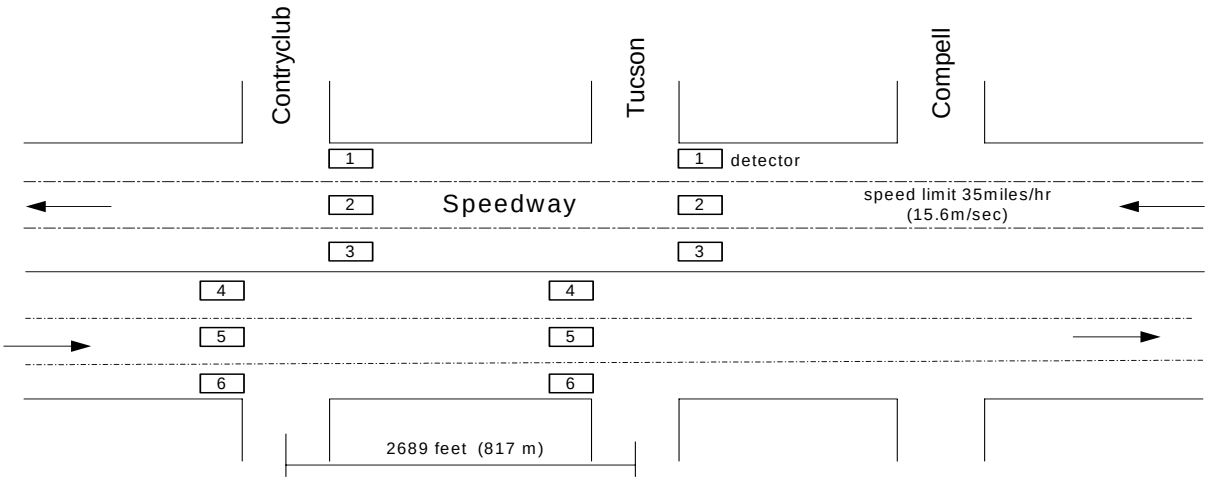


FIGURE 6 Single loop detector locations in Speedway Ave at Tucson.

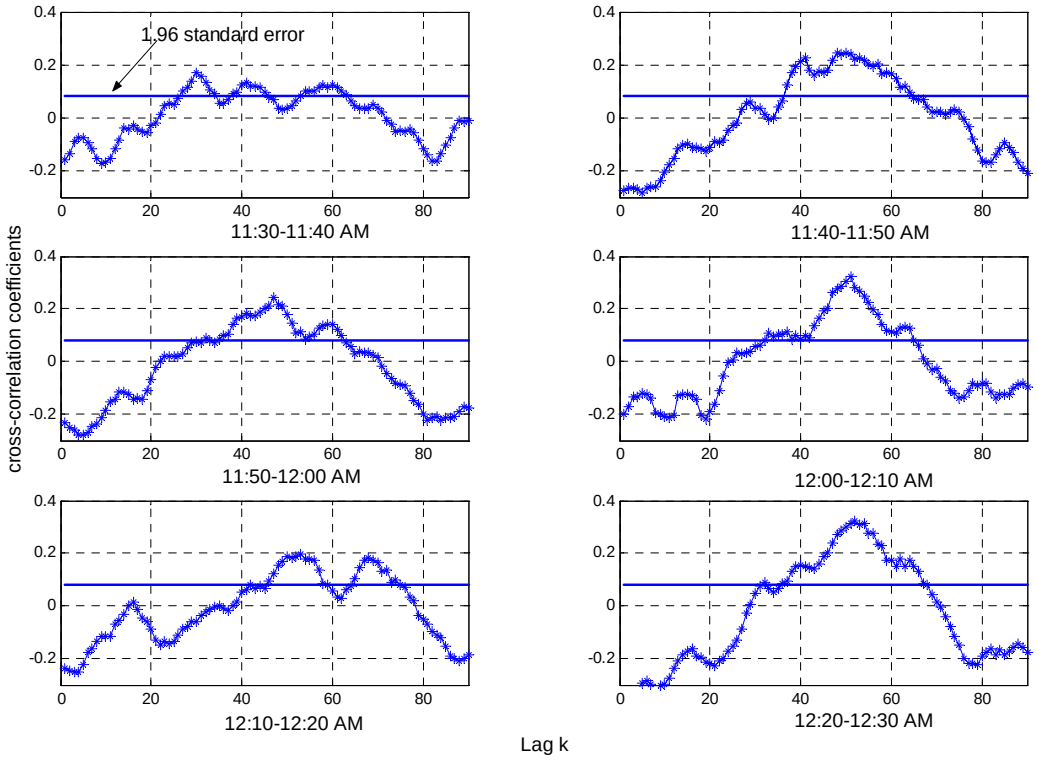


FIGURE 7 Cross-correlation analyses at different time periods.

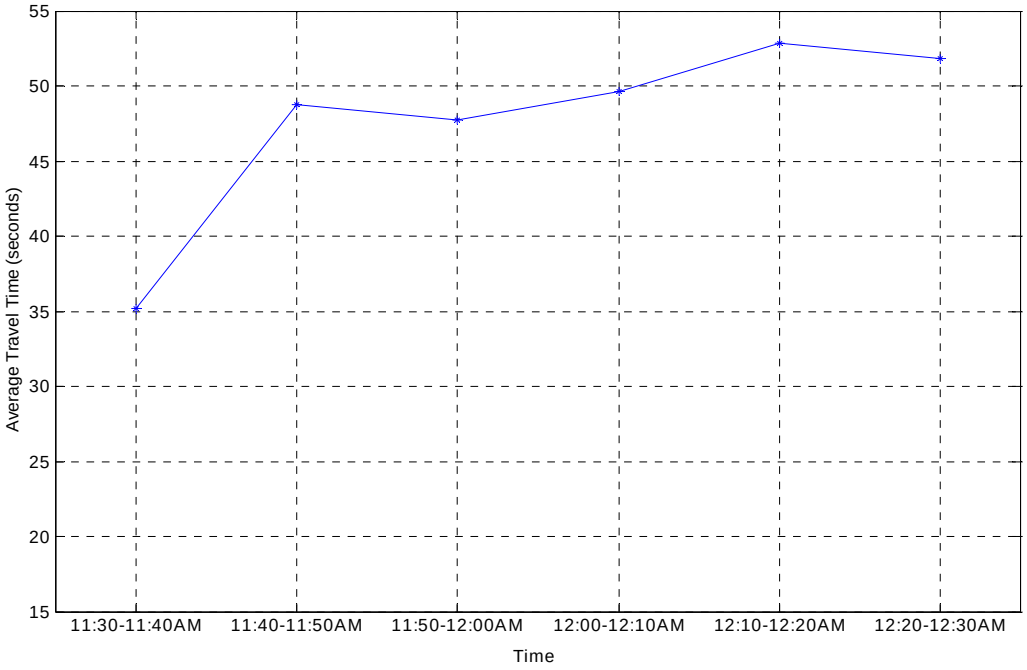


FIGURE 8 Estimated average travel time using the proposed method.