

DEVELOPING TRAVEL TIME ESTIMATION METHODS USING SPARSE GPS DATA

Irum Sanaullah^{1,2}, Mohammed Quddus^{1,*}, and Marcus Enoch¹

¹Transport Studies Group, School of Civil and Building Engineering, Loughborough University, Loughborough, Leicestershire, LE11 3TU, United Kingdom.

²Department of Transportation Engineering and Management, University of Engineering and Technology, Lahore, 54890, Pakistan.

*corresponding author. Email address: m.a.quddus@lboro.ac.uk; Telephone: +44 (0)1509 228545.

Abstract: Existing methods of estimating travel time from GPS data are not able to simultaneously take account of the issues related to uncertainties associated with GPS and spatial road network data. Moreover, they typically depend upon high frequency data sources from specialist data providers which can be expensive and are not always readily available. The study reported here therefore sought to better estimate travel time using ‘readily available’ vehicle trajectory data from moving sensors such as buses, taxis and logistical vehicles equipped with GPS in ‘near’ real-time.

To do this, accurate locations of vehicles on a link were first map-matched to reduce the positioning errors associated with GPS and digital road maps. Two mathematical methods were then developed to estimate link travel times from map-matched GPS fixes, vehicle speeds and network connectivity information with a special focus on sampling frequencies, vehicle penetration rates and time window lengths. GPS data from Interstate I-880 (California, USA) for a total of 73 vehicles over 6 hours were obtained from the UC-

Berkeley's Mobile Century Project, and these were used to evaluate several travel time estimation methods, the results of which were then validated against reference travel time data collected from high resolution video cameras.

The results indicate that vehicle penetration rates, data sampling frequencies, vehicle coverage on the links and time window lengths all influence the accuracy of link travel time estimation. The performance was found to be best in the 5 minute time window length and for a GPS sampling frequency of 60 seconds.

Key words: Travel time estimation, GPS data, Map-matching, vehicle penetration rate, sampling frequency

INTRODUCTION

The estimation of network traffic states such as travel time, traffic flow, speed and density has increasingly become important for travellers and traffic operators (Jiang et al. 2006; Li, 2008), and travel time is a particularly important input for assessing and monitoring the efficiency and performance of a transport network and its associated traffic management systems. For instance, travel time is the primary input to route guidance that provides information to travellers regarding the most suitable route depending on the traveller's requirements, preferences and network states (Liao and Chen, 2015). Simply, a longer travel time will lead to a higher travel cost, and travellers would be able to avoid this situation if reliable pre-trip and en-route travel time information was provided. Therefore, accurate estimation of travel time has become a matter of high importance in recent years (Li, 2008).

Unfortunately, existing methods of estimating travel time from GPS data are not able to simultaneously take account of the issues related to uncertainties associated with GPS and spatial road network data, e.g. sampling frequency, vehicle coverage on the network, time window length and vehicle penetration rate. Moreover, they also tend to depend upon higher frequency data sources from specialist data providers and hence can be expensive and not always readily available. The study reported here therefore sought to better estimate travel time using 'readily available' vehicle trajectory data from moving sensors such as buses, taxis and logistical vehicles equipped with GPS in 'near' real-time.

The paper is structured as follows: the next section reviews the available literature, followed by a brief description of the data used in this study. The section on methodology describes two methods which are used to estimate link travel time using GPS and road network data. The results section then provides the analysis and comparison with the reference travel time data. Finally, the paper ends with conclusions.

LITERATURE

For travel time estimation it is essential to first determine the accurate location of vehicles travelling on road segments (Zhang and Gao, 2008). A range of different map-matching algorithms has emerged in the literature over the last decade. Based on the techniques used in the map-matching process, these map-matching algorithms can be categorised as geometric, topological, probabilistic and artificial intelligence (Xi et al. 2007; Jagadeesh et al. 2004). A geometric MM algorithm is the simplest approach to find a road segment which is the closest to the vehicle trajectory. A topological approach uses the information of the link's geometry as well as link connectivity in the map-matching process. A probabilistic algorithm uses statistical models (i.e. based on probability theory) and considers all the links as the candidate links which fall within the defined error region around a positioning point. A fuzzy logic-based map-matching algorithm is an artificial intelligence technique that uses knowledge-based fuzzy rules formulated from the input and is able to take into account uncertainty associated with the map-matching process. This type of map-matching algorithm is suitable for dense urban road network in which the identification of the correct road segment among the candidate segments is more difficult (Syed and Cannon, 2004). Quddus et al. (2007) presented a detailed description of all existing map matching algorithms from simple (Geometric and probabilistic) to complex approaches (e.g. Fuzzy logic, Extended Kalman Filter, Dempster-Shafer's theory, Bayesian theory, Belief theory and Hybrid Map-matching).

For the development of an efficient traffic management system on freeways and arterials, the accurate travel time estimation is essential, and therefore there are number of methods which have been proposed in literature to estimate the travel times under these conditions (Petty et al. 1998; Krishnan, 2008). Initially for freeway travel time estimation, techniques such as the 'Sandglass Method' (Takaba et al.1991), 'Simple Stochastic Based Regression Method' (Petty et al. 1998), 'K-Nearest Neighbour Model' (Handley et al. 1998) 'Artificial Neural

Network Approach' (Dharia and Adeli, 2003), and 'Spot-Speed Models' (Zhang, 2006) typically drew on vehicle occupancy and flow data from fixed sensors such as Inductive Loop Detectors (ILD). Similarly, for arterial travel time estimation, loop detector occupancy and mean delay data at intersection were correlated using traffic simulation modelling (Bhaskar et al. 2007; Sisiopiku et al. 1994). In general, extrapolation methods are the simplest methods for travel time estimation, using data from ILDs. These methods are based on the assumption that speed is constant for the small distance between measuring points, and the spot speed and link speed can certainly be related in the case of freeways. However an extrapolation method cannot perform well in conditions of congestion and high traffic flow in urban networks, for example, where signalised or non-signalised intersections cause delays which can make the travel time estimation more challenging (Li et al. 2006; Bhaskar et al. 2007).

More recent studies imply that data from fixed sensors may not be sufficient to reliably estimate travel time due to limited coverage area, low data accuracy and high implementation cost (Jiang et al. 2006; Izadpanah et al. 2011), and consequently the focus now is on using moving sensors to provide high accuracy speed (i.e. speeds of travelling vehicles), positioning and timing data to estimate travel time. For travel time estimation on arterials, the travel speed and time method (Yoo et al. 2005), Two-fluid model (Xiang et al. 2007) and Artificial Neural Network model (Zheng and Van Zuylen, 2013) were developed using GPS data from probe vehicles or cell phones.

However, techniques relying on moving sensor data too, have problems that need addressing. Specifically, these relate firstly to the fact that positioning data from GPS have errors, especially in urban areas, and so some form of map-matching technique needs to be applied to match raw positioning fixes on to the road network so as to reliably estimate link travel time. This is challenging because while most current map-matching methods are suitable for high frequency GPS positioning data (i.e. data with 1 second intervals), they may not be

appropriate for lower frequency data (e.g. data with 30 second or 60 second intervals). Unfortunately though, many moving sensors from which data is generally more readily available (such as from bus, taxi, or logistics operating companies for example), only retain low frequency data so as to reduce the cost of data storage and transmission. In addition, any map-matching method to be employed in travel time estimation needs to operate quickly and efficiently so that a large number of GPS data can be processed within a short enough period of time. Secondly, once the map-matching has been completed, the next issue is to accurately estimate link travel time for links, particularly where there is little data available and where the quality of the data which does exist is relatively poor.

In proposing solutions to these issues, Hellings et al. (2008) suggested a method to solve the problem of travel time allocation on individual links (one of the steps necessary to estimate the average travel time on link), which was evaluated using data from a simulated network and compared with the benchmark method. However, although the method was evaluated for different data sampling frequencies (i.e. 15 seconds, 30 seconds, 60 seconds, 90 seconds and 100 seconds), vehicle penetration rates (i.e. proportion of vehicles equipped with GPS) were not considered. It was suggested to evaluate the proposed method for an actual road network. Next, Zheng and Van Zuylen (2010) presented a comparison of travel time estimation models for urban links using probe vehicle data. Their proposed ANN model was evaluated using the VISSIM simulated urban network. The network was simulated for 65 minutes for each simulation run and the data with a 60 second sampling frequency was used. The result was that the ANN model performed better than the Hellings's model. However, there were still problems. Thus, the GPS data used was insufficient (1 car data), and the projection of GPS data on links (map-matching), vehicle penetration rate and its influence on the accuracy of travel time estimation were not addressed (Zheng and Van Zuylen, 2013).

Hunter et al. (2009) meanwhile proposed an expectation maximization algorithm that identifies the possible paths travelled by probe vehicles and estimates the historical link travel time distribution through the network. The main focus of this study was to find the shortest trajectory taken by vehicle from the four selected shortest trajectories. Developing this, Herring et al. (2010) proposed a probabilistic approach to estimate the arterial travel time distribution using sparse probe vehicles data. The model was presented for traffic estimation using the measurements of random locations at random times. Data from 500 taxis with a sampling frequency of 1 minute were used. Here, the factors of varied data sampling frequency and lengths of time window were not considered.

Izadpanah, et al. (2011) studied the real-time travel time prediction for motorways using 1 second data from 8 probe vehicles and shockwave information. Data from probe vehicles provided better accuracy of link travel time estimation compared with data from the detectors. The performance of the model with low frequency data however, was not evaluated. Zheng and Van Zuylen (2013) proposed Artificial Neural Network (ANN) model for the estimation of link travel time

Zhan et al. (2013) developed a model to estimate urban link travel times by using hourly origin destination trip data from taxicabs equipped with GPS receivers. The model was evaluated using one week data (1000 observations in one hour on weekday and 500 observations on a weekend) which included the information about trip distance, trip origin and destination location and path travel time. In practice, the model was found to be capable of effectively estimating average link travel times. However, there were biases related to the route choice model to find the actual path travelled by taxicab; the hourly average estimated values of travel time did not represent the variations in link travel time in one hour; and the non-consideration of the junction delays in the proposed model can lead to the inaccurate estimation of travel time on urban links.

Wang et al. (2014) proposed a real-time model for travel time estimation in a city considering the issues of data scarcity and finding the optimal combination of trajectories for a query path. An attempt was made to address the data infrequency problem by modelling the driver's travel times on road segments with the combination of geospatial, temporal and historical framework of trajectories and dynamic programming solution was devised to find out the best possible linkage of trajectories for a particular path. Two months of GPS data from more than 32,600 taxis with the average sampling rate of 96 sec per point in the time slot of 30 minutes was used for the assessment of proposed method, and was found to perform better than the baseline approaches in finding the correct values for missing data.

Rahmani et al. (2015) used low sampling frequency GPS data of 2 minutes to estimate the path travel time distribution by a non-parametric method. The focus of the study was to determine and correct the biases related to the floating car data for instance incomplete route coverage and traversal of the adjoining road network. Data from 1,500 taxis was used and results were compared with the travel data from ANPR. The proposed method was found to be able to calculate the mean, median and percentiles of the travel time distribution.

In all of these studies the comparative analysis using data with different sampling frequencies in multiple time window lengths was not addressed.

As discussed, the vehicle penetration rate can be defined as the proportion of GPS-equipped vehicles travelling on a road in a given period of time. It is advantageous to have a higher vehicle penetration rate so as to enhance the estimation reliability but this may not always be possible due to issues related to personal privacy and the cost of data collection. It is also difficult to collect and process the data at short intervals of time, so that after a certain point the accuracy of link travel time estimations do not increase significantly with the increase in vehicle penetration rate (Jiang et al. 2006; Xiang et al. 2007). Therefore, an

estimation of optimal vehicle penetration rate is desirable. In this regard, previous studies have shown that 5% of vehicles are sufficient to be used as probe vehicles to obtain the accurate estimates of travel time (Li, 2005; Herrera et al. 2010). Meanwhile, Chen and Chien (2000) used a micro-simulation model to find the minimum number of moving sensors required for travel time estimation and concluded that at least 3% vehicles are needed to provide statistical accuracy in travel time estimation and for heavy flow 12% vehicles are required for accurate estimates. More specifically, when traffic flow is low or high, then the vehicle penetration rate should also be high (Chen and Chien, 2001). They however concluded that the relationship between vehicle penetration rate and prediction error needs further research.

[Table 1]

Table 1: Summary of key studies on travel time estimation methods

Table 1 shows the factors being considered or not considered in different studies in the literature. In order to develop a reliable travel time estimation method, all the factors should be considered and analysed collectively using a large data set. To summarise, the accuracy of travel time estimation using data from moving sensors depends on a range of factors such as vehicle penetration rate (i.e. proportion of vehicles equipped with GPS), coverage of links (i.e. proportion of links with GPS-equipped vehicles) and the process of eliminating positioning errors related to both GPS and road network data. It would also be interesting to investigate the performance of a travel time estimation method for GPS data with different sampling frequencies (such as 1 second, 30 second, 60 second, etc.).

CONCEPTUAL FRAMEWORK

The comparative analysis presented in the above literature review indicates that the factors of different GPS data sampling frequencies, vehicle penetration rates, and time window lengths have not been carried out at the same time in link travel time estimation methods. All of these factors should be considered collectively using data from an actual road network. Consequently, in this paper link travel time estimation algorithms are proposed to estimate link travel time using GPS data consisting of time stamps, latitudes, longitudes, speed and road network data. Specifically, the proposed algorithms are different to existing methods in that these methods are dynamic in nature so as to estimate the travel time for all kinds of possible scenarios related to the inconsistent distribution of data on links (e.g. ranging from the scenario of not having a single GPS point on a link to the scenario of having more than one point from more than one vehicle on the that particular link. In addition to that, the proposed methods incorporate the data from adjacent links and time windows to consider the correlation of travel time in adjacent links (spatially) and over time (temporally) in the link travel time estimation.

Finally the issues with sampling frequency, vehicle penetration rate, time window length, link coverage and map-matched vehicle positions should be considered and analysed, whilst the accuracy of travel time estimation by developed methods should also be validated by using the reference (true) travel time data collected by ANPR cameras, which has been rarely done in literature (Rahmani et al. 2015).

In compiling these elements from the literature review into a conceptual framework, it was anticipated that the performance of the travel time estimation methods would develop as follows:

Sampling frequency: The expectation was that accuracy would exponentially fall as sampling frequencies were progressively reduced from a 1 second sampling frame.

Time Window Length: Increasing the time window length (TWL) was predicted to increase the data coverage on links which was expected to increase the accuracy of results.

Vehicle penetration: Here, results were expected to improve as vehicle penetration increased, up to a level of 5% or so, at which point previous studies suggested that accuracy does not increase significantly.

Model complexity: The expectation was that the more advanced mathematical models developed would outperform the simpler algorithms.

DATA PROCESSING

GPS data collected by the University of California, Berkeley on a section of I-880 as part of the Mobile Century project has been used in this paper. Further details about the project and data collection can be found in Herrera et al. (2010).

Specifically, this paper presents the analysis of data from a total of 73 vehicles equipped with GPS-enabled smart phones (i.e. Nokia N95) which travelled on interstate I-880 California between Decoto Road and Winton Avenue in the Northbound (NB) direction. GPS positioning data points (i.e. longitude and latitude) in WGS84 were converted to the American plane coordinate system (California Zone 3; easting and northing coordinates) using a GIS package. The heading (the direction from the North) at each epoch was calculated using two consecutive easting and northing coordinates. The sampling frequency of the raw GPS data was 3 seconds. Altogether a total of 485,079 GPS observations were processed. In order to see how the performance of a travel time estimation method varies with the sampling frequency, the original GPS dataset was extracted at different sampling frequencies, namely 6, 30, 60 and 120 seconds. Road network data for Interstate I-880 and surroundings were obtained from a location-based service (LBS) provider CloudMade (<http://cloudmade.com>). Reference (true) travel time data from Decoto Road to Winton Avenue recorded by the ANPR cameras were also gathered from the same source so as to estimate the accuracy of the developed travel time estimation methods. A total of 65 links for the time duration of 6 hours were used in the analysis.

METHODOLOGY

Link travel time is defined as the average travel time obtained from all the vehicles that travelled on a link during a particular time window length (e.g. 5 minutes, 10 minutes or 15 minutes). The primary objective here was to develop and evaluate the methods to estimate link travel time using data from GPS-equipped vehicles. Accordingly, the methodology consisted of three main steps: map-matching, link travel time estimation, and validation. These are briefly discussed below.

Step-1: Map-Matching

The first step was to map-match raw GPS positions on the links where the vehicles were travelling. This was achieved by using a map-matching algorithm (Quddus et al. 2003) requiring two types of input data:

1. GPS data including time stamp (t), easting (x), northing (y), speed (v) in m/s and heading (i.e. the vehicle movement direction with respect to the north direction).
2. Road network data consists of link ID number, start node (SN) and end node (EN) with their x and y coordinates (i.e. SN_x , SN_y ; EN_x , EN_y).

Most of the existing map-matching algorithms that use historical map-matched data, link connectivity and turn restriction information at junctions are suitable for 1Hz GPS data (i.e. sampling frequency 1 second) (Quddus et al. 2003). Since the GPS data to be employed in this study contained various (lower) sampling frequencies i.e. 3, 30, 60, 120 seconds, existing map-matching algorithms were not considered to be directly applicable. Therefore, a weight-based topological algorithm developed by Velaga et al. (2009) was employed, which was based on two weights: proximity (the distance for a GPS fix to a candidate link) and heading difference (the difference between the vehicle heading from GPS and link heading) (see Quddus et al. 2003 for further details).

To illustrate this, in Figure 1, P is a vehicle position from a GPS receiver and the corresponding candidate links are BC , BF , AB and BE . Equation (1) was then used to calculate the total weight score (TWS) for each of the candidate links and the link with the highest total weight is chosen as the correct link for the GPS fix. This process was repeated at every GPS epoch. The outputs of each epoch are: map-matched easting and northing coordinates, speed and time stamp.

[Figure 1]

FIGURE 1: Map-matching

$$TWS_i = W_{P_i} + W_{H_i} \quad (1)$$

$$TWS_i = P_w \left[1 - \frac{|D|}{Th} \right] + H_w \cos(\Delta\theta_i) \quad (2)$$

In which,

$$P_w + H_w = 1; \quad 0 < P_w < 1; 0 < H_w < 1$$

Where W_{P_i} represents the weight for proximity and W_{H_i} represents the weight for heading, P_w and H_w , represent the weight coefficients for proximity difference and heading respectively, $\Delta\theta_i$ is the angle between vehicle heading (θ) and link direction (β) with respect to the north, D is the perpendicular distance from a GPS fix to the link and Th represents the threshold for the position error indicating the allowable maximum positioning error in the positioning data from the GPS. Specifically, links which fell within the 100m of a GPS point were selected as the candidate links, meaning that the threshold value for the position error was 100m. The values for both proximity and heading weights were set at 0.5.

Step-2: Link Travel Time Estimation Methods

As discussed in the literature review section, link travel time depends on many factors including the number of map-matched GPS points on the link, speeds of the vehicles, GPS sampling frequencies, time window lengths, vehicle penetration rates and vehicle trajectories. In the first step, links travelled by GPS-equipped vehicles (i.e. link coverage) were identified for the given T_w (i.e. 5 min, 10 min and 15 min). Link-level map-matched points (from Step-1) obtained from those vehicles were then stored. As expected, there was a mixed distribution of map-matched points by link. In representing these, four types of scenario were tested:

- Scenario 1: no map-matched point on a link.
- Scenario 2: only one map-matched point from one vehicle on a link.
- Scenario 3: more than one map-matched point from one vehicle on a link.
- Scenario 4: more than one map-matched point from more than one vehicle on a link.

In this paper two methods were proposed to estimate link travel times from GPS data under different scenarios. Method-1 considered map-matched points from adjacent links while estimating travel time to enhance the link coverage and to reduce the uncertainty in the estimation. Method-2 assimilated link travel times from Method-1 in both spatial and temporal ways. Both of these methods are discussed below:

6.1 *Method-1 Distance and Time Proportion*

Map-matched fixes from step-1 were employed to estimate link travel time. Figure 2 depicts a generic condition in which there are four map-matched points from two different vehicles on link AB. One approach would have been to determine average speed on link AB based on speed measurements (from GPS) associated with these four map-matched points and then use average speed and the length of the link to estimate link travel time. Alternatively, the average speed for each of the vehicles could have been calculated and then employed to estimate two

measurements of travel time. Both of these approaches however would likely result in an incorrect estimate of travel time, especially for an urban link where delays are more common at junctions due to traffic lights. To counter this effect, the solution arrived at created vehicle trajectories by considering not only all map-matched points on link AB but also map-matched points outside link AB (i.e. map-matched points beyond downstream and upstream junctions of link AB). Thus, vehicle trajectories were created by taking into account map-matched points $M_3 - M_6$ as well as $M_1 - M_2$ and $M_7 - M_8$ (see Figure 2). After that the distance and time proportion approach for the vehicle trajectories (beyond and within the link) was utilised to reliably estimate travel time for portions of the link without any map-matched points. The distance and time proportion method for travel time estimation for V_i is explained in Figure 2. As can be seen in Figure 2, the travel time for link AB consisted of the cumulative travel time for the three portions of the link (see Equation 3). The distance and time proportion method was then applied for the first and last portion of the link (TT_1 and TT_2) as shown in Equations 3(a) and (b). The travel time for the middle portion (*i.e.* $\Delta t_{i+2,i+1}$) was obtained from the time differences at epoch t_{i+2} and t_{i+1} as shown in Equation (3).

[Figure 2]

FIGURE 2: Distance and time proportion travel time estimation

$$TT_{AB} = \Delta t_{i+2,i+1} + \sum_{i=1}^2 TT_i \quad (3)$$

$$TT_1 = (t_{i+1} - t_i) \times d_{i+1} / (d_i + d_{i+1}) \quad 3(a)$$

$$TT_2 = (t_{i+3} - t_{i+2}) \times d_{i+3} / (d_{i+3} + d_{i+4}) \quad 3(b)$$

However, it should be noted that the above equations were applicable only when there was at least one map-matched point from the adjacent links is available. All possible conditions are depicted in Figure 3 and fully explained below.

[Figure 3]

FIGURE 3: Summary of distance and time proportion algorithms

Model 1: When no map-matched fixes from the adjacent links are available in T_w

The generic form of travel time estimation by method-1 for model-1 is presented in Figure 4.

Scenario 1: For the first scenario when no map-matched point existed on a link in a particular T_w , then speed limit data of the link was considered as being the ‘average speed’ to use in estimating the travel time. Another option is to employ link-based historical travel time that would be more accurate than that of using the speed limit.

Scenario 2: For the second scenario, when one map-matched point was available on a link, then the distance between the two nodes of link (AB) and vehicle speed (v) at epoch (t) given by the GPS receiver was used for link travel time estimation as shown in Figure 3(b).

Scenario 3: For the third scenario when there were more than one map-matched points from one vehicle, the average speed ($\bar{v}$) of that vehicle was used to estimate the travel time for the first and last portion of a link as shown in Figure 3(c). The link travel time therefore consisted of the travel time for each of the three portions of the link, denoted by TT_1 , TT_2 and $\Delta t_{i+3,1}$ as shown in Figure 3(c) and Equations 4(a) and (b).

$$TT_1 = (|x_i^m - x_1|^2 - |y_i^m - y_1|^2)^{1/2} / \bar{v} \quad 4(a)$$

$$TT_2 = (|x_{i+1} - x_{i+3}^m|^2 - |y_{i+1} - y_{i+3}^m|^2)^{1/2} / \bar{v} \quad 4(b)$$

Where x_i , y_i and x_i^m , y_i^m were the easting and northing of start node and first map-matched point on a given link respectively.

Scenario 4: For the fourth scenario when map-matched points were from more than one vehicle passing on the link, the travel time was estimated for each vehicle using average speeds and time stamp data as explained in Scenario 3 and then the average travel time for the link was calculated as shown in Figure 3(d).

Where TT_v represents travel time for a vehicle and n is the number of vehicles travelled on a link.

[Figure 4]

FIGURE 4: Flow chart of Model 1 algorithm of Method 1

Model 2: When map-matched data were only available from one adjacent link (either the succeeding or the preceding link) in T_w

In Model 2, map-matched points on the adjacent link were considered along with the map-matched points on the link for which the link travel time was being estimated (see Figure 2). This data included the map-matched positions, time stamps and speed data both for both links. For these models the distance and time proportion method was used to estimate travel time for a given link under different scenarios as shown in Figure 3 (e), (f), (g), (h) and explained below.

Scenario 1: same as Scenario-1 under Model-1

Scenario-2: travel time for the first portion of the link (TT_l) was estimated using the distance and time portion method between the last map-matched point of preceding link and the first map-matched point on a given link as shown in Figure 3(f). While the travel time for the remaining part of the link was estimated using vehicle speed at epoch (t_{i+1}) and distance

between the map-matched point and the end node of link. Where t_i and t_{i+1} are the time epochs and d_i and d_{i+1} represent the distance of map-matched points on a preceding and given links respectively.

Scenario 3: since there was more than one map-matched point on a link, the total travel time therefore consisted of the aggregate of the travel times for each of the three portions of the link as shown in Figure 3(g). For the first portion, travel time (TT_1) was estimated using the distance and time proportion method. For the last portion of the link, travel time (TT_2) was estimated using average speed as shown in Figure 3(g).

Scenario 4: there were more than one map-matched points from more than one vehicle travelling on the given link as shown in Figure 3(h). Link travel time here was estimated as Scenario-3 for each vehicle separately and then the average travel time was calculated based on travel time experienced by all vehicles.

The above description was also applicable for the case of a succeeding link with one map-matched point as shown in the Figure 3 (i), (j), (k), (l).

Model 3: when map-matched data were available from both adjacent links in T_w

Model 3 was based on the condition whereby map-matched points on both preceding and following links were available for a link in the same time window length. Therefore, the time distance proportion method was applied both for the first and last portion of given link as described below and shown in Figure 2 and Figure 3(n).

Scenario 1: same as Scenario 1 Model 1.

Scenario 2: the map-matched data from following and preceding links were available along with the data on the current link. Therefore travel time for the first and last portion of the given link (TT_1 and TT_2) was estimated based on the position and time data at epoch (t_i), (t_{i+1}) and (t_{i+2}). The application of time and distance proportion method is shown in Figure 3(n).

Scenario 3: the travel time estimation approach used has been explained for the example of vehicle 1 (V_1) in Figure 2 and Equation (3).

Scenario 4: the link travel time was estimated as discussed above for Scenario 3. That is, individually for all vehicles and then the average travel time of link was calculated based on the average travel time experienced by all of the vehicles on the given link as shown in Figure 3(p).

6.2 Method-2 Spatial and Temporal Moving Average

It was reasonable to assume that the link travel time among adjacent links may be correlated with each other for a given time window. For instance, the link travel time for link (AB) may be correlated with the link travel time for both links (DA) and (BC), especially for the same time window length (see Figure 5). It was also rational to assume that link travel time of a link may be correlated over time. Especially, link travel time for link (AB) at time window (t) may be correlated with travel time at time window ($t-1$) and ($t+1$). Therefore, the estimation of link travel time could potentially be enhanced by link travel time input values from the adjacent links for the same time window (i.e. spatial component) and from the adjacent time windows for the same link (i.e. temporal component). Since vehicle trajectory data from GPS were used in estimating travel time, for the temporal component it was required to integrate the current travel times of a given link with those of the previous and following time windows. For the spatial component, the lengths and travel times of the connected links needed to be used. Therefore, the link travel time estimation based on spatial and temporal component was derived as follows:

[Figure 5]

FIGURE 5: Link travel time using the spatial and temporal connectivity

Link travel time for link AB:

(a) Spatial component – since link travel time for link AB was likely correlated with that of link DA and BC, the weighted average of link travel time of these three links was judged to provide a good estimate of link travel time for AB.

$$T_{AB}^s = (T_{DA} * L_1 + T_{AB} * L_2 + T_{BC} * L_3) / (L_1 + L_2 + L_3) \quad (5)$$

(b) Temporal component – average link travel time for AB over n th consecutive time windows:

$$T_{AB}^t = \frac{1}{n} \sum_{k=-n}^n (T_{AB})_{t-k} \quad (6)$$

Where n was the total number of time windows considered and $n=1$.

This suggests that travel times for the given link in the current time window can be estimated using the travel times in both the preceding and succeeding time windows. By taking $n=1$ in Equation (6) it is confirmed that link travel time is estimated by integrating the travel times of the current time window (t), the preceding time window ($t-1$) and the succeeding time window ($t+1$). This ensures that the method is applicable to real-time (Method-1) and near real-time (Method-2) with a delay of 5-15 minutes. For $n>1$, the method would therefore be suitable for an offline application.

Total link travel time for link AB was obtained as the weighted average of spatial and temporal components as follows:

$$TT_{AB} = \alpha(T_{AB}^s) + \beta(T_{AB}^t) \quad (7)$$

In which $\alpha + \beta = 1$ and $0 < \alpha < 1; 0 < \beta < 1$

Where T_{AB}^s, T_{AB}^t represent the spatial and temporal components of travel time for link AB, L_1, L_2, L_3 represent the length of links DA, AB and BC respectively, and α, β are the weight coefficients for spatial and temporal components respectively. The values of these two weight coefficients ($\alpha = 0.1$ and $\beta = 0.9$) were determined based on an empirical analysis. This suggests that the effect of spatial correlation in link travel time estimation was in fact much weaker than initially hypothesised. It is worthwhile to note that the values of the two weight coefficients α and β can also be obtained from link-based reference travel time data if such data are available.

RESULTS

Evaluation Indicators

For the purpose of performance evaluation two widely used statistics such as Mean Absolute error (MAE) and Mean Absolute percentage error (MAPE) were used (3):

$$MAE = \frac{1}{N} \sum_{i=1}^N |TT_{r,i} - TT_{e,i}| \quad (8)$$

$$MAPE = 100 * \frac{1}{N} \sum_{i=1}^N \left| \frac{TT_{r,i} - TT_{e,i}}{TT_{r,i}} \right| \quad (9)$$

Where, $TT_{r,i}$ is the average reference travel time from probe vehicles observed during time window length i , $TT_{e,i}$ is the mean estimated travel time during time window length i and N is the total number of time windows.

The reference travel time data recorded by high resolution ANPR video cameras provided the exact travel times of individual vehicles through license plate re-identification. A total of 4,126 vehicles were matched during the experiment from 11:00 to 17:00. From the travel times of these vehicles, the average reference travel time was calculated for 5, 10 and 15 minute time window lengths. The average link travel times estimated by the two developed algorithms for different sampling frequency GPS data sets (i.e. 3, 30, 60, 120 seconds) in three time window lengths (i.e. 5, 10 and 15 minutes) were compared with the average reference link travel time from 11:00 to 17:00 as shown in Figure 6. The solid line represents the average link reference travel time while the dotted lines show the link travel time estimated by the developed algorithms for each of the different sampling frequency GPS data sets (i.e. 3, 30, 60, 120 seconds).

[Figure 6]

FIGURE 6: Comparison of link travel time estimates for 5 minutes time window length

The overall accuracy of both methods for the 5 minute time window length was found to be better relative to the other time window lengths (i.e. 10 minute and 15 minute) due to the higher vehicle penetration rate (5-8%) and higher vehicle coverage on the test route as shown in Figure 7. The mean vehicle penetration rate ranged between 5-8% for the 5 minute time window length, 4-5% for the 10 minute and 3-4% for the 15 minute time window length. The correlation between the estimated link travel times by the two methods and the reference link travel times for different sampling frequency data sets (i.e. 3, 30, 60, 120 seconds) is shown in Figure 8. For the data sets with sampling frequencies of 3, 30 and 60 seconds, the distance and time proportion method (i.e. Method 1) provided good correlations between the estimated and the reference travel time (i.e. the R^2 was more than 0.8). However for the data set with a sampling frequency of 120 seconds, the travel time estimated by Method 1 differed from that of the reference travel time significantly as shown in Figure 8(g). Meanwhile the spatial and temporal moving average method (i.e. Method 2) performed better than the distance and time proportion method which can be inferred from Figure 8(b) (d) (f) and (g). Specifically it showed a high correlation with the reference travel time (the R^2 was more than 0.90 for the 30 and 60 second data sets) and performed better even for the 120 second data set as compared to the distance and time proportion method.

[Figure 7]

FIGURE 7: Change in vehicle penetration rate with time window lengths

[Figure 8]

FIGURE 8: Correlation between reference travel time and estimated travel times

The performance of the link travel time estimation results was also evaluated in terms of MAPE and MAE as shown in Table 1. The MAPE and MAE were small for both methods when the sampling frequency of data sets was 60 and 120 seconds, though interestingly the prediction was less accurate for the data sets with higher 3 and 30 second sampling frequencies for both methods. Overall then, the spatial and temporal moving average method performed better than the distance and time proportion method by reducing the MAPE from 8.56% and 11.12% to 7.55% and 8.43% for 60 and 120 second data sets respectively.

[Table 2]

Table 2: Performance Comparison between Distance and Time Proportion Method and Spatial and Temporal Moving Average Method (5 minutes T_w)

DISCUSSION AND CONCLUSION

This paper identified a range of factors that influence the estimation of link travel time from GPS data of sub optimal quality (i.e. lower frequencies) of the type that is most readily available. This includes: GPS sampling frequency, vehicle penetration rate and length of time window. The paper then argued that one needs to control for all of these factors simultaneously in order to reliably estimate link travel time from vehicles equipped with GPS. Two mathematical methods were therefore developed. Method 1 was based on the number of map-matched points and connectivity of links and Method 2 was based on spatial and temporal travel time components of the link. This paper then examined the performance of these two travel time estimation algorithms using real-world data from a section of Highway I-880.

Specifically, in practice the elements in the conceptual framework did not perform quite as expected in several regards.

Thus, while the expectation for sampling frequency was that accuracy would exponentially fall as sampling frequencies were progressively reduced from a 1 second sampling frequency, in practice both methods performed better for the lower frequency GPS data sets (i.e. 60 and 120 seconds) as compared to the higher frequency data sets (i.e. 3 and 30 seconds). At first glance this seems surprising in the first instance as one would expect that higher frequency data would lead to a better estimate of travel time. However, whilst this may be true for traffic in free-flowing conditions, in a congested network a higher GPS sampling frequency means that there are many GPS observations travelling at low speeds which could somehow impact on the accuracy of travel time predictions.

Such an explanation would tally with the traffic conditions experienced on the test route, whereby an accident resulted in slow traffic between 10.00 and 11.50 on the day for which the

data were obtained, whilst the route additionally suffered heavy congestion from 15.20 onwards (see Mazare et al. 2012 for details). In this situation, it is conceivable that for the 3 and 30 second frequency tests a higher proportion of map-matched points were sampled travelling at low speeds thereby leading to the over-estimation of travel time and hence to errors in the travel time estimation prediction. By contrast for the lower frequency cases the GPS data may have been masked, resulting in smaller errors in travel time. Moreover, this effect may have been exacerbated because the travel time from Decoto Road to Winton Avenue (i.e. a section of the test route I-880) was calculated by summing up the link travel times of each of the 65 individual links between this O-D pair, travel time was overestimated due to traffic congestion. Finally, the data set with a sampling frequency of 60 seconds was found to be optimal with the MAE of 1.15 min (MAPE 8.56%) and 1.04min (MAPE 7.55%) by the first and second link travel time estimation methods respectively.

Similarly, whilst increasing the TWL was predicted to increase the link coverage and finally the accuracy of travel time estimation, it was found that actually the 5 min TWL had better results than the 10 and 15min TWLs. One reason for this could be that the high vehicle penetration rate (5-8%) was observed in 5 min TWL as compared to 4-5% and 3-4% vehicle penetration rate in 10 min and 15 min TWL respectively. In future the impact of time window length can be further investigated while keeping the constant rate of penetration in all time window lengths.

Regarding vehicle penetration, results were expected to improve as vehicle penetration increased, up to a level of 5% or so, it was found that in the time window length of 5 min when the vehicle penetration rate was 5-8%, the accuracy of results improved.

Finally, although both methods performed well for travel time estimation, spatial and temporal moving average method (Method 2) was found to be the best performing model for

low sampling frequency GPS data. This appears to have been because in Method 2, the link coverage was improved due to the data infusion from adjacent links and from previous and following time windows.

Moreover, since the computation times for the travel time estimation methods was less than 3 minutes for the 3 second data set, but less than 1 minute for the 30, 60 and 120 second data sets, the developed methods would perhaps be more suitable for real-time ITS applications.

From the results it is concluded that the accuracy of the developed method depends upon establishing and maintaining a consistent and optimal vehicle penetration rate for a particular time window length, thus ensuring a reasonable coverage of the network in terms of the number of map-matched points on each link.

Based on the above discussions, the following recommendations are made for future research:

- (a) The weight coefficients for spatial and temporal components derived empirically can be further checked for different road network data to find the optimal values using an optimisation technique;
- (b) The impact of lengths of time windows can be further investigated by keeping the vehicle penetration rate constant;
- (c) In future the analysis with urban road network data can be carried out;
- (d) The combination of data from different sources, for example, Bluetooth, smart phones, mobile sensors and inductive loop detectors data can be used to analyse the accuracy of travel time.

REFERENCES

- Bhaskar, A., Chung, E., De Mouzon, O., & Dumont, A. G. (2007). Methodology for Travel Time Estimation on a Signalised Arterial. In *Proceedings of the Young Researchers Seminar* (No. LAVOC-CONF-2008-015).
- Chen, M., & Chien, S. I. J. (2000). Determining the number of probe vehicles for freeway travel time estimation by microscopic simulation. *Transportation Research Record: Journal of the Transportation Research Board*, **1719**(1), 61-68.
- Chen, M., & Chien, S. I. J. (2001). Dynamic freeway travel-time prediction with probe vehicle data: Link based versus path based. *Transportation Research Record*, **1768**(1), 157-161.
- Dharia, A., & Adeli, H. (2003). Neural network model for rapid forecasting of freeway link travel time. *Engineering Applications of Artificial Intelligence*, **16**(7), 607-613.
- Handley, S., Langley, P., & Rauscher, F. A. (1998). Learning to predict the duration of an automobile trip. *Proceedings of 1998 International Conference on KDD and Data Mining*, New York City. 27-31 August, 219–223.
- Hellinga, B., Izadpanah, P., Takada, H., & Fu, L. (2008). Decomposing travel times measured by probe-based traffic monitoring systems to individual road segments. *Transportation Research Part C: Emerging Technologies*, **16**(6), 768-782.
- Herrera, J. C., Work, D. B., Herring, R., Ban, X. J., Jacobson, Q., & Bayen, A. M. (2010). Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment. *Transportation Research Part C: Emerging Technologies*, **18**(4), 568-583.
- Herring, R., A. Hofleitner, P. Abbeel, & A. Bayen (2010). Estimating arterial traffic conditions using sparse probe data. Paper presented at the *13th International IEEE Conference on Intelligent Transportation Systems*, Madeira Island, Portugal, 19–22 September.
- Hunter, T., R. Herring, P. Abbeel, & A. Bayen (2009). Path and travel time inference from GPS probe vehicle data. Paper presented at the Workshop on Analysing Networks and

Learning with Graphs, *23rd Annual Conference on Neural Information Processing Systems*, Whistler, BC, 11 December.

Jagadeesh, G., Srikanthan, T. & Zhang, X. (2004). A map matching method for GPS based real-time vehicle location. *The Journal of Navigation*, vol. 57, no. 03, pp. 429-440.

Izadpanah, P., Hellinga, B., & Fu, L. (2011). Real-time freeway travel time prediction using vehicle trajectory data. (Paper 11-2826) Presented at the *Transportation Research Board, 90th Annual Meeting*, Washington, D.C., 23-27 January.

Jiang, G., Gang, L., & Cai, Z. (2006). Impact of probe vehicles sample size on link travel time estimation. Paper presented at the *Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE*, 505-509, Toronto, Ontario, 17-20 September.

Krishnan, R. K. (2008). *Travel time estimation and forecasting on urban roads* (PhD Thesis). Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London.a

Li, R., Rose, G. & Sarvi, M. (2006). Evaluation of speed-based travel time estimation models. *Journal of Transportation Engineering*, vol. 132, no. 7, pp. 540-547.

Li, W., Chuanjiu, W., Xiaorong, S., & Yuezuo, F. (2005). Probe vehicle sampling for real-time traffic data collection. *Proceedings of IEEE Intelligent Transportation Systems*. Vienna, 13-15 September, 222-224.

Li, Y. (2008). Short-term prediction of motorway travel time using ANPR and loop data. *Journal of Forecasting*, **27**(6), 507-517.

Liao, C. H., & Chen, C. W. (2015). Use of Advanced Traveler Information Systems for Route Choice: Interpretation Based on a Bayesian Model. *Journal of Intelligent Transportation Systems*, (ahead-of-print), 1-10.

Mazaré, P. E., Tossavainen, O. P., Bayen, A., & Work, D. B. (2012). Trade-offs between inductive loops and GPS probe vehicles for travel time estimation: A mobile century case study. (Paper 12-2746) Paper presented at the *Transportation Research Board 91st Annual Meeting*, Washington, D.C., 22-26 January.

- Petty, K. F., Bickel, P., Ostland, M., Rice, J., Schoenberg, F., Jiang, J., & Ritov, Y. A. (1998). Accurate estimation of travel times from single-loop detectors. *Transportation Research Part A: Policy and Practice*, 32(1), 1-17.
- Quddus, M.A., Ochieng, W.Y. & Noland, R.B. (2007). Current map-matching algorithms for transport applications: State-of-the art and future research directions, *Transportation Research Part C*, 15(5), 312-328.
- Quddus, M. A., Ochieng, W. Y., Zhao, L., & Noland, R. B. (2003). A general map matching algorithm for transport telematics applications. *GPS Solutions*, 7(3), 157-167.
- Rahmani, M., Jenelius, E., & Koutsopoulos, H. N. (2015). Non-parametric estimation of route travel time distributions from low-frequency floating car data. *Transportation Research Part C: Emerging Technologies*.
- Robinson, S., & Polak, J. W. (2005). Modeling urban link travel time with inductive loop detector data by using the k-NN method. *Transportation Research Record*, 1935, 47-56.
- Syed, S. & Cannon, M. (2004). Fuzzy logic based-map matching algorithm for vehicle navigation system in urban canyons. *ION National Technical Meeting, San Diego, CA*.
- Sisiopiku, V. P., & Roupail, N. M. (1994). Toward the use of detector output for arterial link travel time estimation: a literature review. *Transportation Research Record*, (1457).
- Takaba, S., Morita, T., Hada, T., Usami, T., & Yamaguchi, M. (1991). Estimation and measurement of travel time by vehicle detectors and license plate readers. *Proceedings of the Vehicle Navigation and Information Systems Conference*, Dearborn, MI, 20-23 October, 2, 257-267.
- Velaga, N. R., Quddus, M. A., & Bristow, A. L. (2009). Developing an enhanced weight-based topological map-matching algorithm for intelligent transport systems. *Transportation Research Part C: Emerging Technologies*, 17(6), 672-683.
- Wang, Y., Zheng, Y., & Xue, Y. (2014). Travel time estimation of a path using sparse trajectories. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 25-34). ACM.

Xi, L., Liu, Q., Li, M. & Liu, Z. (2007). Map Matching Algorithm and Its Application. *ISKE proceedings*, Chengdu, China.

Xiang, Z., Jian, R., Jiancheng, W., & Changqiao, S. (2007). Network-wide performance assessment of urban traffic based on probe vehicle data. *Proceedings of the Intelligent Transportation Systems Conference, ITSC 2007. IEEE*, Seattle, WA, 30 September – 3 October, 442-447.

Yoo, B.S., Kang, S.P., Park, C.H. (2005). Travel time estimation using mobile data. *Proceedings of Eastern Asia Society for Transportation Studies*, Tokyo, Japan, vol. 5, pp. 1533–1547.

Zhan, X., Hasan, S., Ukkusuri, S. V., & Kanga, C. (2013). Urban link travel time estimation using large-scale taxi data with partial information. *Transportation Research Part C: Emerging Technologies*, 33, 37-49.

Zhang, W. (2006). *Freeway Travel Time Estimation Based on Spot Speed Measurements*, PhD Dissertation. Virginia Polytechnic Institute and State University, Virginia.

Zhang, Y. & Gao, Y. (2008). An improved map matching algorithm for intelligent transportation system. *IEEE International Conference on Industrial Technology*, pp.1-5.

Zheng, F., & van Zuylen, H. (2010). Comparison of urban link travel time estimation models based on probe vehicle data. *Proceedings of the 7th International Conference on Traffic and Transportation Studies, ICTTS 2010*, Kunming, 3-5 August, 615-626.

Zheng, F., & Van Zuylen, H. (2013). Urban link travel time estimation based on sparse probe vehicle data. *Transportation Research Part C: Emerging Technologies*, 31, 145-157.

LIST OF TABLES

Table 1: Summary of key studies on travel time estimation methods

Table 2: Performance Comparison between Distance and Time Proportion Method and Spatial and Temporal Moving Average Method (5 minutes Tw)

Table 1: Summary of key studies on travel time estimation methods

References	Data	Factors				
		Different Data Sampling Frequencies	Vehicle Penetration Rate	Time Window Lengths	Map-Matching	ANPR Data For Validation
Rahmani et al. (2015)	1500 taxis (2 min GPS data)	×	×	×	✓	✓
Wang et al. (2014)	32,00 taxis (96 sec GPS) data	×	×	×	✓	×
Zhan et al. (2013)	OD trip GPS taxicab data (1 week)	×	×	×	✓	×
Zheng and Van Zuylen, (2013)	1 probe vehicle (60 sec GPS data)	×	×	×	×	×
Izadpanah, et al. (2011)	8 probe vehicles 1sec GPS Data	×	×	×	×	×
Hellinga et al. (2008)	Simulation Data	✓	×	×	✓	×

× not considered ✓ considered

Table 2: Performance Comparison between Distance and Time Proportion Method and Spatial and Temporal Moving Average Method (5 minutes T_w)

Data sets	Distance and Time proportion		Spatial and temporal moving average	
	MAPE (%)	MAE(min)	MAPE (%)	MAE(min)
3 sec	13.40	1.56	15.60	1.86
30 sec	11.76	1.65	13.39	1.90
60 sec	8.56	1.15	7.55	1.04
120 sec	11.12	1.45	8.43	1.07

LIST OF FIGURES

FIGURE 1: Map-matching

FIGURE 2: Distance and time proportion travel time estimation

FIGURE 3: Summary of distance and time proportion algorithm

FIGURE 4: Flow chart of Model 1 algorithm of Method 1

FIGURE 5: Link travel time using the spatial and temporal connectivity

FIGURE 6: Comparison of link travel time estimates for 5 minutes time window length

FIGURE 7: Change in vehicle penetration rate with time window lengths

FIGURE 8: Correlation between reference travel time and estimated travel times

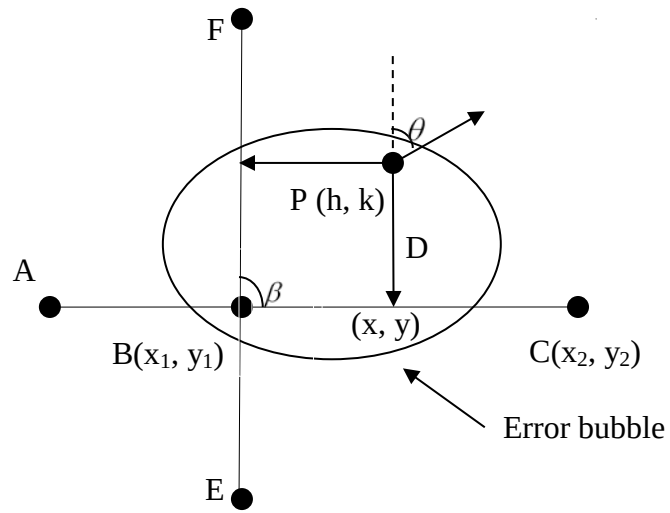


FIGURE 1: Map-matching

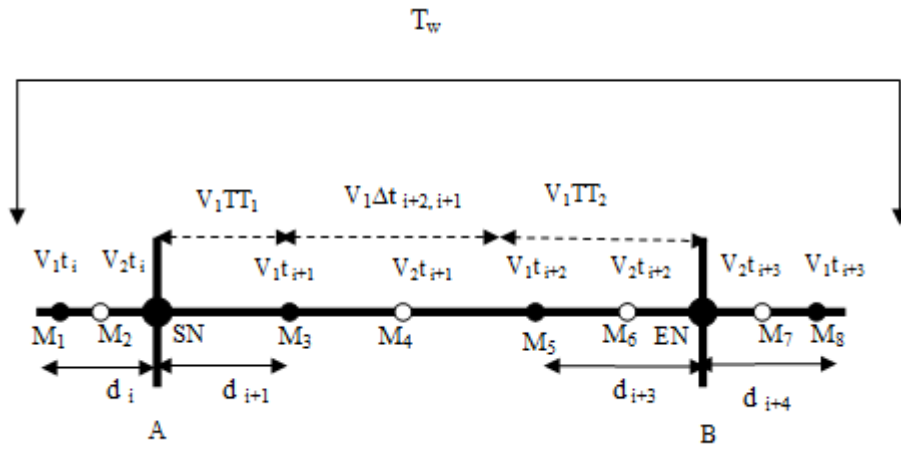


FIGURE 2: Distance and time proportion travel time estimation

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Model-1	(a)	(b)	(c)	(d)
	$TT_L = d/v_s$	$TT_L = d/v_t$	$TT_L = \Delta t_{i+3,i} + \sum_{i=1}^2 TT_i$	$TT_L = \frac{1}{n} \sum_{v=1}^n TT_v$
Model-2	(e)	(f)	(g)	(h)
	$TT_L = d/v_s$	$TT_1 = (t_{i+1} - t_i) \times d_{i+1} / (d_i + d_{i+1})$ $TT_2 = d_{i+2} / v_{t_i}$	$TT_1 = (t_{i+1} - t_i) \times d_{i+1} / (d_i + d_{i+1})$ $TT_2 = d_{i+3} / \bar{v}$	$TT_L = \frac{1}{n} \sum_{v=1}^n TT_v$
Model-3	(i)	(j)	(k)	(l)
	$TT_L = d/v_s$	$TT_1 = d_i / v_{t_i}$ $TT_2 = (t_{i+1} - t_i) \times d_{i+1} / (d_{i+1} + d_{i+2})$	$TT_1 = d_i / \bar{v}$ $TT_2 = (t_{i+2} - t_{i+1}) \times d_{i+2} / (d_{i+2} + d_{i+3})$	$TT_L = \frac{1}{n} \sum_{v=1}^n TT_v$
Model-3	(m)	(n)	(o)	(p)
	$TT_L = d/v_s$	$TT_1 = (t_{i+1} - t_i) \times d_{i+1} / (d_i + d_{i+1})$ $TT_2 = (t_{i+2} - t_{i+1}) \times d_{i+2} / (d_{i+2} + d_{i+3})$	$TT_L = \Delta t_{i+4,i+1} + \sum_{i=1}^2 TT_i$	$TT_L = \frac{1}{n} \sum_{v=1}^n TT_v$

FIGURE 3: Summary of distance and time proportion algorithm

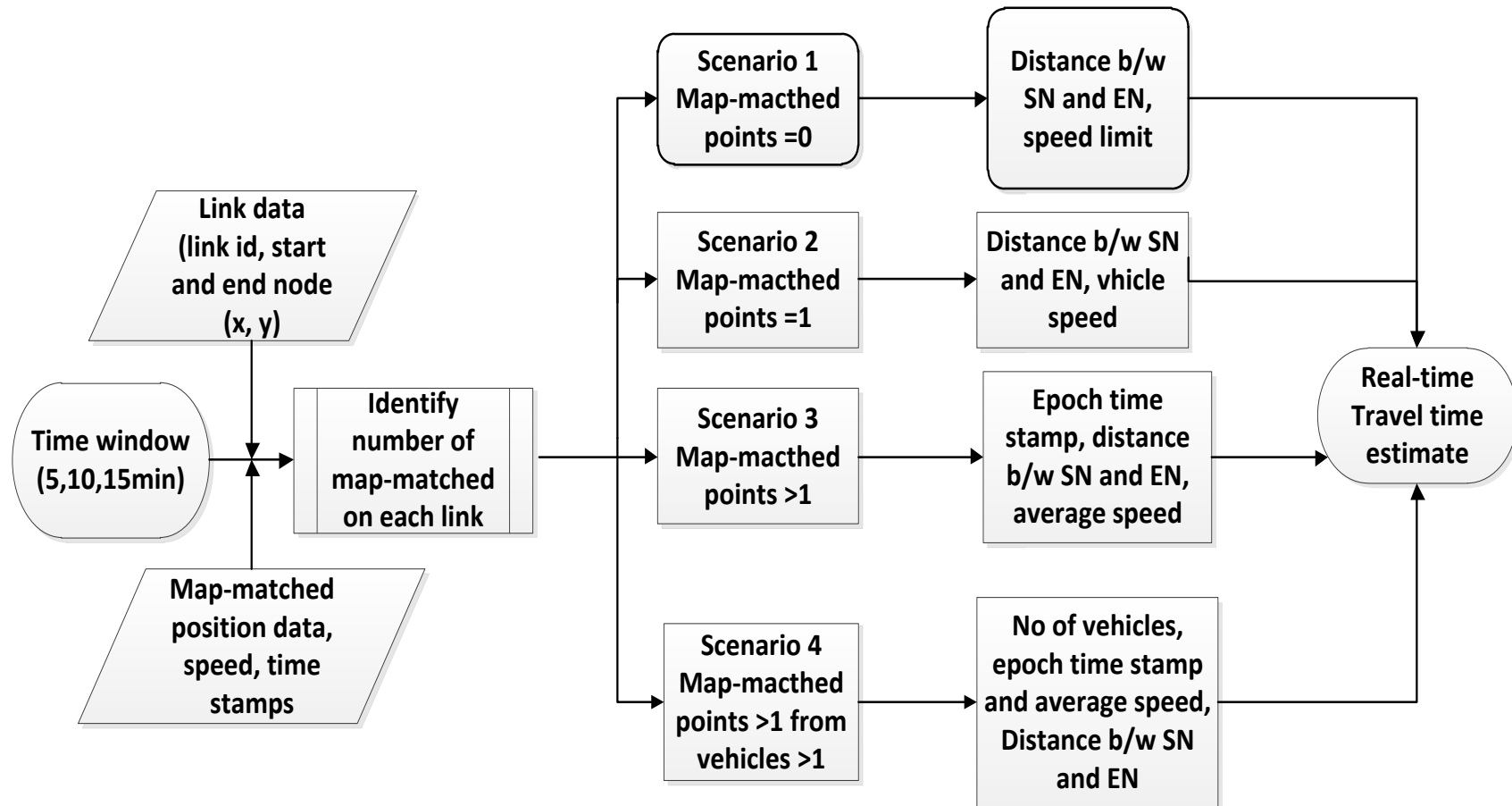


FIGURE 4: Flow chart of Model 1 algorithm of Method 1

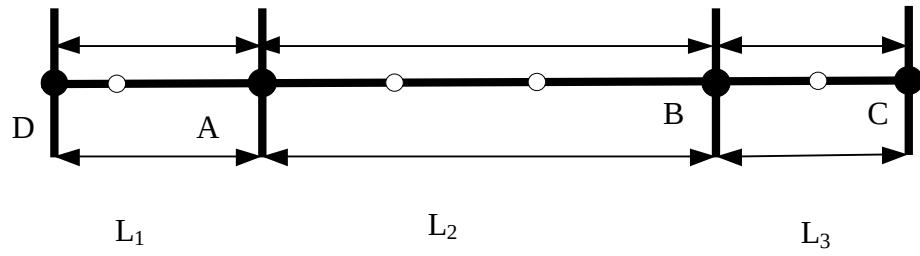
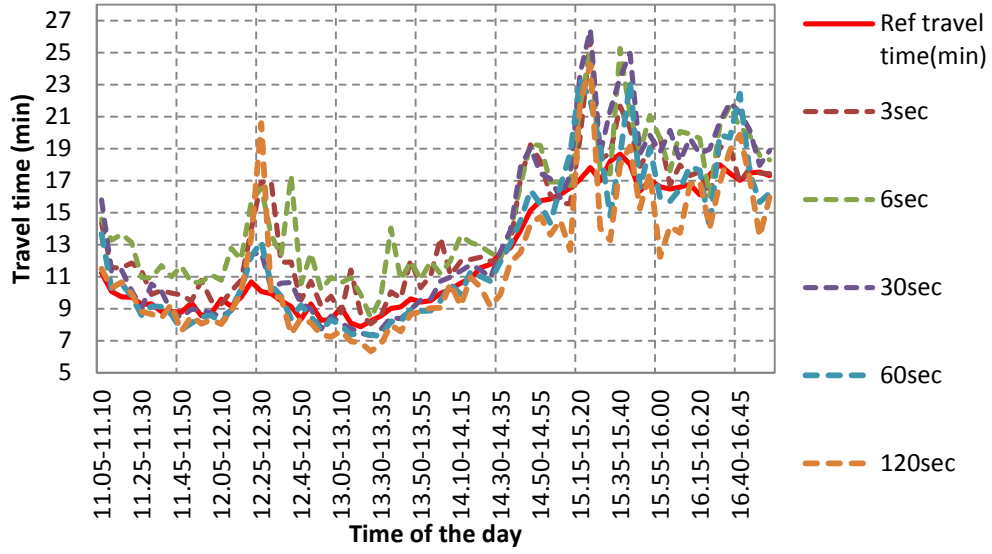
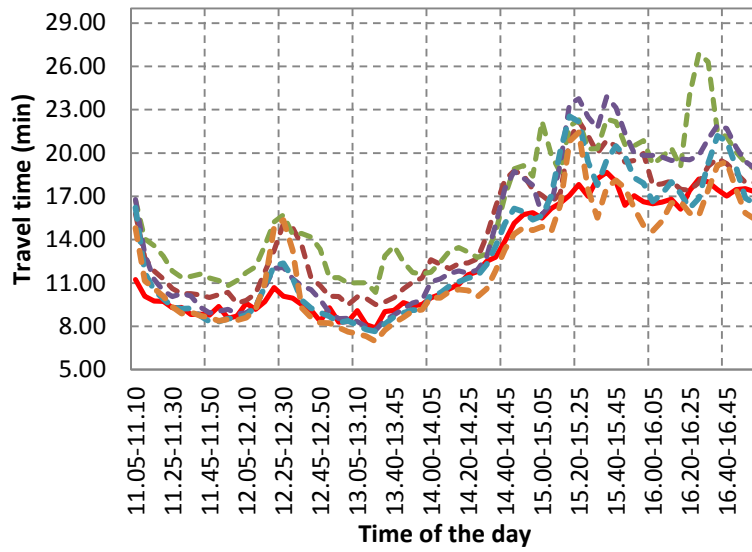


FIGURE 5: Link travel time using the spatial and temporal connectivity



(a) Distance time proportion method



(b) Spatial and temporal moving average method

FIGURE 6: Comparison of link travel time estimates for 5 minutes time window length

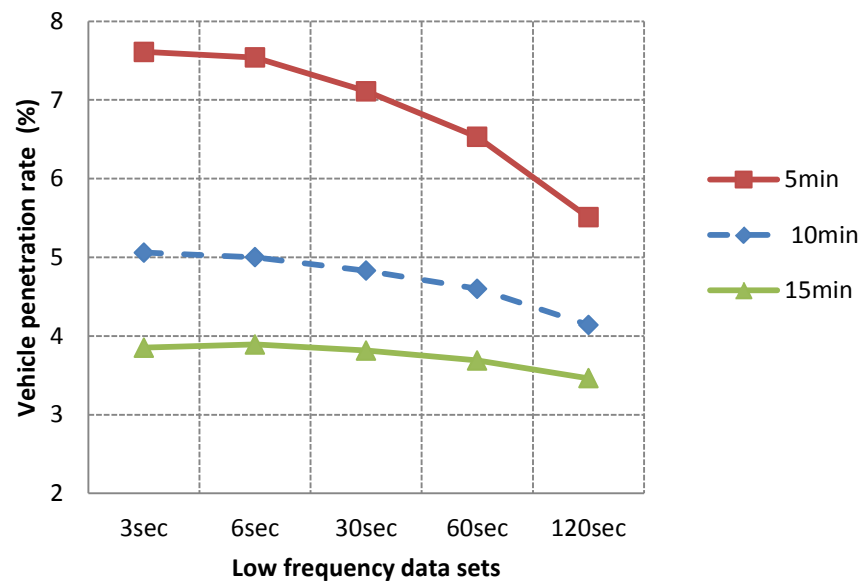
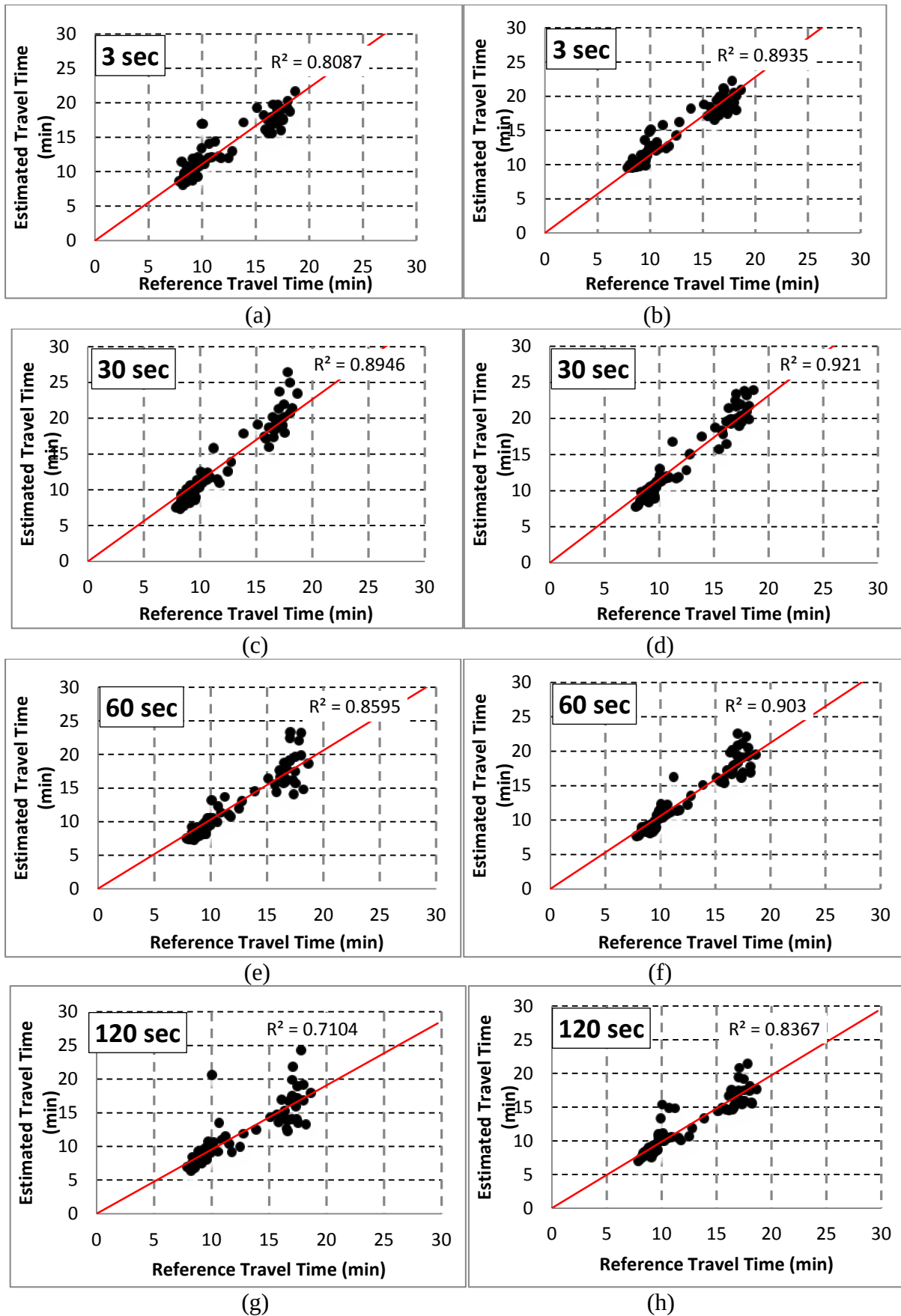


FIGURE 7: Change in vehicle penetration rate with time window lengths



Distance and time proportion method

Spatial and temporal moving average method

FIGURE 8: Correlation between reference travel time and estimated travel times