

## Accepted Manuscript

Utilizing data mining techniques to predict expected freeway travel time from experienced travel time

Hasan M. Moonam, Xiao Qin, Jun Zhang

PII: S0378-4754(18)30026-0

DOI: <https://doi.org/10.1016/j.matcom.2018.01.006>

Reference: MATCOM 4538

To appear in: *Mathematics and Computers in Simulation*

Received date: 9 October 2017

Revised date: 27 January 2018

Accepted date: 29 January 2018

Please cite this article as: H.M. Moonam, X. Qin, J. Zhang, Utilizing data mining techniques to predict expected freeway travel time from experienced travel time, *Math. Comput. Simulation* (2018), <https://doi.org/10.1016/j.matcom.2018.01.006>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



## Utilizing Data Mining Techniques to Predict Expected Freeway Travel Time from Experienced Travel Time

Hasan M. Moonam<sup>a, 1</sup>, Xiao Qin<sup>a, \*</sup>, Jun Zhang<sup>b, 2</sup>

<sup>a</sup>Civil and Environmental Engineering, University of Wisconsin-Milwaukee,  
3200 N Cramer St., Milwaukee, WI 53211-3314, United States

<sup>b</sup>Electrical Engineering & Computer Science, University of Wisconsin-Milwaukee

### ABSTRACT

1 As the most important real-time traveler information, travel time can be either experienced or  
2 expected (i.e. to be experienced). When a vehicle completes a trip, the travel time refers to the  
3 experienced travel time. In contrast, when a vehicle starts its journey, the travel time is unknown  
4 but can be predicted, which is the expected travel time. Although the experienced travel time is  
5 termed as the real-time travel time, a traveler may encounter a somewhat different travel time  
6 (from expected travel time) due to the changing traffic conditions. Therefore, expected travel  
7 time needs to be predicted. In this study, the expected travel time was predicted from the  
8 experienced travel time using the data mining techniques such as k-nearest neighbor (k-NN),  
9 least squares regression boosting (LSBoost) and Kalman filter (KF) methods. After comparing  
10 the performances of KF to corresponding modeling techniques from both link and corridor  
11 perspectives, it is concluded that the KF method offers superior prediction accuracy in a link-  
12 based model. Moreover, the effect of different noise assumptions was examined and it is found  
13 that the steady noise computed from the full-dataset had the most accurate prediction. A data  
14 processing algorithm, which processed more than a hundred million records reliably and  
15 efficiently was also introduced.

16 **Keywords:** experienced and expected travel time, arrival and departure time based travel time,  
17 travel time prediction, data mining, Kalman filter, modeling Kalman filter noise, K-nearest  
18 neighbor method, Boosting, LSBoost.

### 19 1. Introduction

20 Travel time is an important component of Advanced Traveler Information Systems (ATIS),  
21 as it is a key factor for travelers who are faced with non-recurring congestion (Khattak et al.,  
22 1996). Aside from measuring transportation system performance, travel time has been used to  
23 predict future travel time and traffic state, which help the traffic operations room in versatile  
24 ways. Amongst all available techniques, Bluetooth has emerged as one of the fastest growing  
25 data collection technologies whose market share is continuing to rise, mainly due to its cost  
26 effectiveness (Blogg et al., 2010, Moghaddam and Hellinga, 2013). Bluetooth is a probe-based

\* Corresponding author. Tel: +1 (414) 229 7399

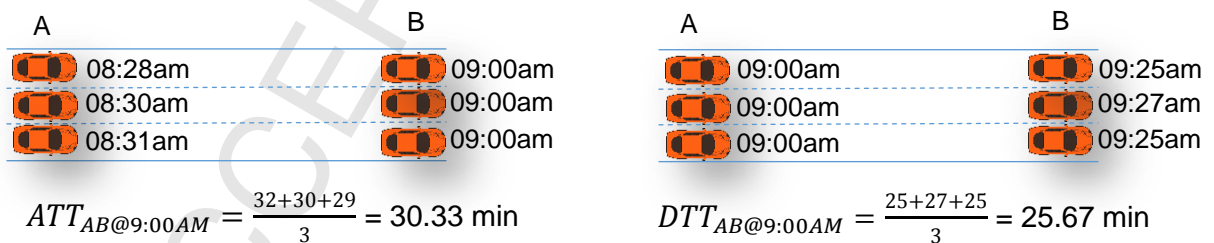
E-mail addresses: [hmoonam@uwm.edu](mailto:hmoonam@uwm.edu) (H. M. Moonam), [qinx@uwm.edu](mailto:qinx@uwm.edu) (X. Qin), [junzhang@uwm.edu](mailto:junzhang@uwm.edu) (J. Zhang)

<sup>1</sup>Tel: +1 (414) 229 7399, <sup>2</sup>Tel: (414) 229 4246

(Puckett and Vickich, 2010) Automatic Vehicle Identification (AVI) technique used for collecting travel time data. Each Bluetooth device contains a unique electronic identifier known as a Media Access Control (MAC) address/identifier. Devices with MAC addresses that are in range can be logged as long as a simple antenna is mounted adjacent to the roadway. Travel time and corresponding traffic speed can be estimated by the timestamp difference between two consecutive stations with matching MAC addresses (Bachmann et al., 2013).

Travel time is measured by the time elapsed when a traveler moves between two distinct spatial positions (Carrion and Levinson, 2012). Over a decade, various studies have attempted to define travel time estimation (e.g. instantaneous travel time, experienced travel time (Xiao et al., 2014), and predicted travel time (Bhaskar et al., 2011)), but the definitions lack clarity, which has created confusion and inconsistency in data collection and analysis (Bhaskar et al., 2011, Toppen and Wunderlich, 2003). Recently, one study made clear the distinction between arrival time-based link travel time and departure time-based link travel time (Kim et al., 2009).

Arrival time-based link travel time (ATT) and departure time-based link travel time (DTT) have two different estimation algorithms. ATT refers to the travel time associated with arrival at the destination, while DTT refers to the travel time associated with departure from the origin. In practice, ATT is the experienced travel time ( $t_{expn}$ ) that is calculated using arrival and departure times of the vehicles when both are available. On the other hand, DTT is the expected travel time ( $t_{expt}$ ) that is predicted at the time of departure when arrival time is unavailable. To get ATT or  $t_{expn}$  and DTT or  $t_{expt}$ , assume that two vehicles ( $V_1$  and  $V_2$ ) start at 8:30am and 9:00am from point A, respectively. If the assumed clock time is now 9:00am and the first vehicle ( $V_1$ ) has just arrived at point B. Then, the arrival time based travel time for link AB,  $ATT_{AB@9:00AM} = 30mins$  (based on the experienced travel time of  $V_1$ ). The departure time based link travel time,  $DTT_{AB@9:00AM} = unknown$ . If the arrival time for  $V_2$  at point B could be predicted (say, 9:25am), the departure time based travel time would be,  $DTT_{AB@9:00AM} = 25mins$  (based on the expected travel time of  $V_2$ ). Since the DTT at 9:00am is unavailable until a later time, i.e., until the  $V_2$  travels the link AB, it is understandable that the DTT at 9:00am requires a prediction of travel time. Fig. 1 illustrates the concept of estimating ATT or  $t_{expn}$  and DTT or  $t_{expt}$  for a route considering multiple vehicles:



**Fig. 1.** Estimating ATT and DTT of link AB at 9am.

Practitioners usually treat ATT (or  $t_{expn}$ ) as the travel time due to the lack of available DTT (or  $t_{expt}$ ); however, this reported ATT is one-step (step interval = travel time) earlier than the actual travel time to be experienced (DTT) by drivers. Although ATT and DTT differ slightly in a free-flow condition, the difference can sharply escalate at the onset and end of traffic congestion. ATT usually lags behind DTT during a transition of traffic state, and the difference

1 starts to decrease when the traffic state becomes stable. Information on travel time during a  
2 transitional state, as opposed to a stable state, is more important to travelers. Ideally, the travel  
3 time should be the predicted travel time that will be experienced by a traveler, or DTT. This  
4 predicted travel time also helps ensure proper and proactive operations and management of  
5 traffic in a network. Unfortunately, few studies have distinguished between DTT and ATT or  
6 attempted to estimate DTT (Kim et al., 2009).

7 The overarching goal of this study is to develop a comprehensive model for short-term  
8 freeway travel time prediction using Bluetooth data. A dynamic filtering algorithm was proposed  
9 to accurately estimate ATT and thus, reliably predict DTT. An efficient computer algorithm was  
10 developed to process, refine, and integrate a massive Bluetooth dataset, which filtered the travel  
11 time. Finally, prediction algorithms were examined to predict DTT from real-time ATT, and the  
12 better performance of the proposed prediction was observed.

## 14 2. Literature review

15 Travel time data are subject to outliers. The main purpose of outlier detection algorithms is to  
16 detect extreme travel times that result from sampling bias. Fixed-range outlier filtering methods  
17 are not suitable for travel time filtering due to local travel time turbulences, especially when they  
18 occur at the onset or end of congestion. To avoid imposing arbitrary fixed-bound, researchers  
19 have introduced moving average speed based lower and upper-bound (Haghani et al., 2010) and  
20 data-driven real-time adaptive-bound methods (Dion and Rakha, 2006). (Dion and Rakha, 2006)  
21 incorporated a few simple yet significant alterations in their proposed adaptive method, which  
22 offers an alternative to conventional algorithms like percentile, deviation, and traditional  
23 (modified) z- or t-statistical test (Liu, 2008, Clark et al., 2002). The main alteration includes  
24 expanding the data validity window when three consecutive observations fall either above or  
25 below (same side) the window. While this key adjustment helps capture sudden changes in travel  
26 time trends, it is prone to the inclusion of extreme outliers, and therefore compromises the  
27 accuracy of travel time estimation. In response to (Dion and Rakha, 2006) method, (Moghaddam  
28 and Hellinga, 2014) proposed a proactive method that uses a pattern recognition model, which  
29 showed superior performance. But the author acknowledged that the performance of outlier  
30 detection algorithms cannot be objectively quantified when the algorithms are applied to field  
31 data. Appropriate estimation of travel time is possible only when an effective outlier filter is  
32 used. Many studies have examined accurate estimation of travel time in a real-time fashion (Dion  
33 and Rakha, 2006, Skabardonis and Geroliminis, 2005, Moghaddam and Hellinga, 2014, Rice and  
34 Van Zwet, 2004). In most cases, the sophistication of filtering algorithms to maximize the  
35 accuracy led to a certain level of complexity in real-time applications. Therefore, a simplified  
36 version of these proposed algorithms is preferred.

37 Broadly, travel time prediction methods can be classified into the classical approach (Oda,  
38 1990) including statistical (Rice and Van Zwet, 2004) and time series models (Al-Deek et al.,  
39 1998, Hamed et al., 1995), and the data mining approach (Vlahogianni et al., 2014, Zheng and  
40 Van Zuylen, 2013, Zhang et al., 2014, Wu et al., 2004, Myung et al., 2011). Due to the instability  
41 of traffic states, most classical approaches have shown to be incapable of better prediction,  
42 especially with regard to structured and unstructured data (Vlahogianni et al., 2014). **Therefore,**  
43 **advanced data mining methods have become popular to predict travel time. As such, (Jenelius**

1 and Koutsopoulos, 2017) proposed a multivariate probabilistic principal component analysis  
2 method that predicts travel time based on the expected distribution of link travel times. The  
3 method provides superior results to the k-NN method. (Zhong et al., 2017) introduced an online  
4 travel time prediction system without sacrificing computational efficiency. The system adopts  
5 functional principal component analysis framework and utilizes historical and real-time travel  
6 time data to predict link travel time. (Zhang et al., 2015) applied a two-component generalized  
7 autoregressive conditional heteroskedasticity (GARCH) model that captures trend and seasonal  
8 components to improve prediction results. (Fei et al., 2011) employed a Bayesian inference-  
9 based dynamic linear model (DLM) to predict online short-term travel time. (Sumalee et al.,  
10 2013) estimated dynamic stochastic journey time distribution and predict travel time based on  
11 stochastic cell transmission model. (Zhan et al., 2013) extracted travel time data from origin-  
12 destination dataset by minimizing the least squared error between the observed and expected  
13 path travel times. (Zou et al., 2014) exploited a space-time diurnal method to predict travel time  
14 that considers spatial and temporal correlation and diurnal pattern of travel times. In addition to  
15 vehicle trajectory-based methods, (Celikoglu, 2013) used flow model to predict travel time that  
16 not only increases accuracy but also reduces computational complexity. Since data mining  
17 approaches fit easily with massive datasets, researchers have also applied neural networks  
18 (Zheng and Van Zuylen, 2013), fuzzy and evolutionary techniques (Zhang et al., 2014), support  
19 vector regression (Wu et al., 2004), and the k-nearest-neighbor (k-NN) (Myung et al., 2011)  
20 model to directly or indirectly predict travel time. According to Myung et al., the use of non-  
21 representative samples to train the artificial neural-network (ANN) model may lead to a non-  
22 negligible error in prediction (Myung et al., 2011). These data mining approaches require  
23 representative samples (Smith et al., 2002) and sometimes suffer from a lack of interpretability  
24 and transferability. On the other hand, simple methods (e.g. instantaneous, historic average and  
25 clustering over specific days) exhibit low accuracy (Van Hinsbergen et al., 2007). Although  
26 LSBoost, a data mining technique, has been widely used (Jiang, 2001a, Jiang, 2001b,  
27 Barutcuoğlu and Alpaydın, 2003, Jiao et al., 2006, Darwish, 2013) in different fields for many  
28 years, only recently being popular in transportation studies such as perception reaction time  
29 (Elhenawy et al., 2017), congestion duration (Ghosh et al., 2016), freight flow (Moscoso- López  
30 et al., 2016) and emission (Oduro et al., 2015) prediction.

31 KF, an optimal recursive data processing algorithm, has been widely used with various  
32 modifications (e.g. adaptive KF (Guo et al., 2014) and extended KF (Liu et al., 2006)) in several  
33 studies including those on travel time prediction (Chien and Kuchipudi, 2003, Nanthawichit et  
34 al., 2003, Yang, 2005, Chen and Chien, 2001). KF incorporates all information that can be  
35 provided and processes all available measurements to estimate the current value of the variables  
36 of interest (Maybeck, 1990). The KF method has two components: process/system/state-space  
37 and measurement/observation. Nanthawichit et al. formed their state-space equation by declaring  
38 traffic density and space mean speed as state variables, developed their observation equation by  
39 declaring traffic volumes and spot speeds as observation variables (Nanthawichit et al., 2003).  
40 Chen and Chien used travel time as the input variable in both of these equations; previous step  
41 travel time was multiplied by a transition matrix to obtain the state update equation (Chen and  
42 Chien, 2001). A similar study was conducted using field data (Chien and Kuchipudi, 2003)  
43 rather than simulated data (Chen and Chien, 2001). Despite promising results, these studies lack  
44 details about sources of process and measurement (variables') values.

1 Researchers (Chen and Chien, 2001, Chien and Kuchipudi, 2003, Nanthawichit et al., 2003,  
2 Yang, 2005, Chen and Rakha, 2014) have modeled the state-space as a linear system to which  
3 KF was applied. The KF model uses both a priori (derived by state-space/process) and the  
4 observation of the same timestamp to get a posterior by using the update equation. Within this  
5 procedure, the model combines all available observations and prior knowledge of a system in a  
6 way that statistically minimizes errors (Maybeck, 1990). The model can be directly applied to  
7 signal processing and control systems (when observations are available) in order to filter out the  
8 noise. KF has also been widely applied to time series forecasting with a state-space model  
9 (Durbin and Koopman, 2012, Hamilton, 1994, Harvey, 1990) where the update equation  
10 (Commandeur and Koopman, 2007, Durbin and Koopman, 2012) clearly states a time latency;  
11 the next step is predicted using the observations from the current step. Hence, predicting DTT  
12 from ATT is more advantageous over a single source time series data models.

### 13 **3. Data preparation and reduction**

14 Bluetooth data contains three variables: the MAC ID of the detector, MAC IDs of the  
15 detected devices, and the detection timestamp. In spite of a simple data format, a complex  
16 processing algorithm is required to produce the final dataset from the source, which stores the  
17 entire network data in a single table. For a logged MAC ID, recorded timestamps at two  
18 consecutive stations are processed to estimate travel time and the corresponding traffic speed.  
19 The detailed description of the processing algorithm is beyond the scope of this paper; therefore,  
20 a brief description is included below with a limited description of the data characteristics.

21 The selected study area consists of a 62.8-mile long route, or approximately 47.5 miles on I-  
22 90 and the remaining on the Beltline Highway in Madison, Wisconsin. The route is equipped  
23 with 41 unequally spaced Bluetooth stations, resulting in 40 links. The first 21 links are on I-90,  
24 the 22<sup>nd</sup> link is on both corridors, and the remaining links are on the Beltline. The spacing varies  
25 from 1.3-3.4 miles on I-90 and 0.4-1.3 miles on the Beltline Highway. Forty-seven days' worth  
26 of data (11/16/2015-01/01/2016) containing more than 100 million records was collected from  
27 traffic in both directions. Half of the records were from outside the study-area. Each station of  
28 the one hundred stations selected captured around one million records for 47 days, or 67,680  
29 minutes. However, a large portion (approx. three-fourths) of the data are either corrupted or  
30 contaminated due to multiple detections and unsuccessful detections (i.e. not detected in two  
31 consecutive stations). [Fig. 2](#) shows the complete procedure of data processing.

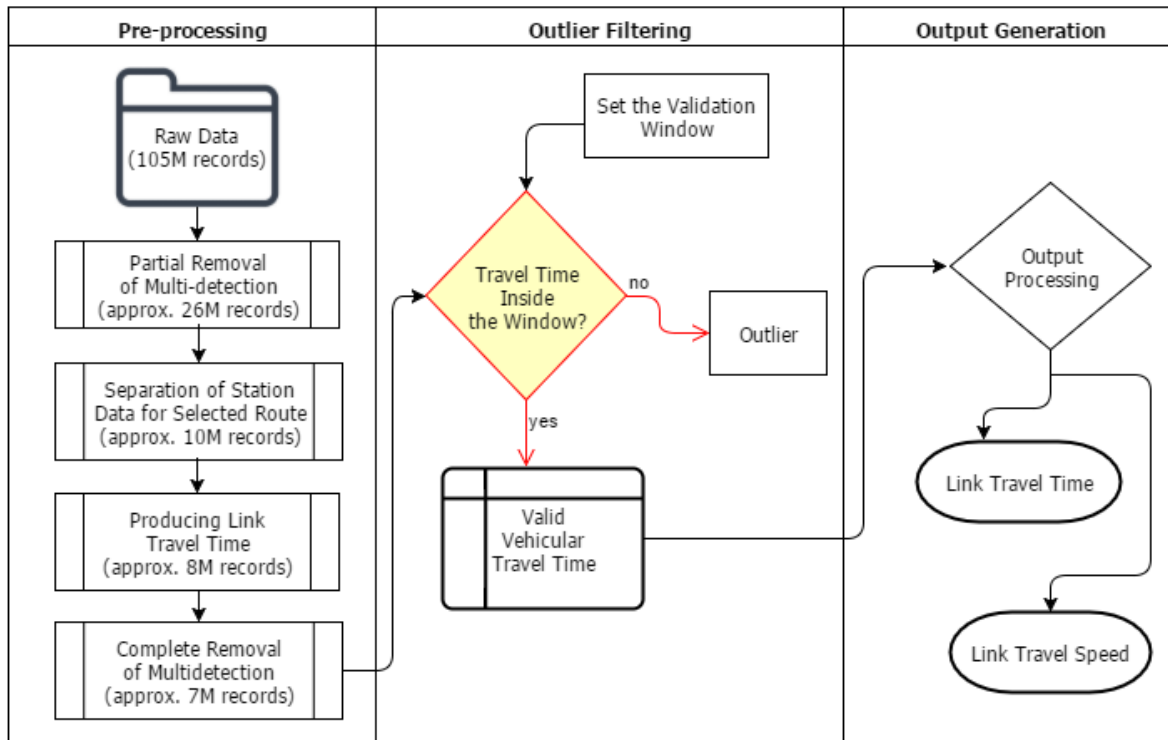


Fig. 2. Data processing procedures.

A Bluetooth station usually detects a Bluetooth device in its range more than once. The number of such detections can increase significantly due to planned or unplanned slowing down/stopping of vehicles. A general inspection of the dataset revealed that such detections usually vary two to four times. Oracle queries helped clean up the multi-detection, resulting in the total number of records decreasing from 105 million to 26 million. The data was then separated by each station for the selected routes, further reducing the records to 10 million. Unsuccessful detections were automatically ignored due to the vehicle's detection timestamps from two adjacent stations. Travel times were calculated. Next, the reduced dataset of 8 million samples was processed through a robust Java-based pre-processing module that investigated each record individually and cleaned all redundant records based on the following principle:

*Two detections of a vehicle at a station ( $STA_1$ ) are valid separate detections if the vehicle is detected at least once at its upstream station ( $STA_2$ ) within the time gap of two detections at  $STA_1$ .*

For example, a vehicle detected on 08:59am, 09:01am and 09:08am at  $STA_1$ , and on 09:04am at  $STA_2$ . Detection on 09:01am is redundant since there is no detection at *upstream station*  $STA_2$  in between 08:59am and 09:01am. In addition, neither 09:01am nor 09:08am is a redundant detection since there is a detection at *upstream station*  $STA_2$  on 09:04am. Note that the Bluetooth stations were capturing both directions of traffic and the vehicle used as an example made a U-turn/return-trip. The pre-processed dataset of 7 million records was further processed to filter outliers. Finally, a Java-based programming module produced the travel time and speed data using the outlier-filtered data. Since travel direction is pertinent to travel time, this study used northbound data.

1

2 **4. Methodology**3 The methodology section details the algorithms for outlier filtering, ATT, DTT, travel speed  
4 estimation, and travel time prediction.

5

## 6 4.1. Outlier Filtering

7 A preset upper and lower boundary helped filter out the outliers. The following equation  
8 defines the lower boundary:

9 
$$tt_{lowr} = tt_{ff} / 2 \quad (1)$$

10 where,  $tt_{lowr}$  and  $tt_{ff}$  stand for lower bound and free flow travel time, respectively. A vehicle  
11 was considered to be an outlier if its speed exceeded more than double the posted speed limit.12 A dynamic validation window works best in outlier filtering (Dion and Rakha, 2006). The  
13 upper boundary is defined by the following equation:

14 
$$tt_{uppr} = tt_e + n \cdot \sigma_e \quad (2)$$

15 where  $tt_{uppr}$ ,  $tt_e$  and  $\sigma_e$  are upper bound, expected travel time, and expected standard deviation  
16 of travel time (samples), respectively.  $n = 1, 2, 3, \dots$ 

17

## 18 4.2. Speed Estimation

19 The following equation estimates the space mean speed of a link AB with length L:

20 
$$v = \frac{L}{\frac{1}{n} \sum_i tt_i} \quad (3)$$

21 where n is the observation count in a defined interval and  $tt_i$  is the travel time of  $i^{\text{th}}$  observation.

22

## 23 4.3. Travel Time Prediction

## 24 4.3.1. K-Nearest Neighbor (k-NN) method

25 Travel time at any timestamp is related to the travel time of its close temporal proximity.  
26 Therefore, DTT is modeled by the nearest ATT.

27 
$$DTT_t = ATT_t + \Delta tt_t \quad (4)$$

28 where  $\Delta tt_t$  = predicted difference of ATT and DTT at a time-step t.29 Travel time difference,  $\Delta tt_t$  is predicted by the distance weighted k-NN method using  
30 historic daily data:

31 
$$\Delta tt_t = \frac{\sum_d w_{d,t} \Delta tt_{d,t}}{\sum_d w_{d,t}} \quad (5)$$



1 where  $\Delta tt_{d,t}$  is the difference between ATT and DTT on a historic day, d at a time-step, t and  
 2  $w_{d,t}$  is the corresponding weight which is the measure of the similarity between two traffic  
 3 patterns: the traffic pattern of the present day and of the historic day, d. This similarity is the  
 4 reciprocal of the measure of the variation between those two traffic patterns. This variation is  
 5 measured by the Euclidian squared distance of two n-dimensional vectors representing the latest  
 6 travel times (ATT of n-steps) from the present day, p and historic day, d:  
 7  $[ATT_t \text{ } ATT_{t-1} \text{ } ATT_{t-2} \text{ } \dots \text{ } \dots \text{ } ATT_{t-n+1}]_p$  and  $[ATT_t \text{ } ATT_{t-1} \text{ } ATT_{t-2} \text{ } \dots \text{ } \dots \text{ } ATT_{t-n+1}]_d$ .  
 8 Since, the n-steps should be the steps that have the most significant impact on the current step t  
 9 to reflect the traffic pattern, the determination of n is a heuristic approach. In this study, travel  
 10 time was predicted for  $n = 1, 2, 3, \dots, 10$ .

11

#### 12 4.3.2. Boosting: Least Square Regression (LSBoost) method

13 LSBoost is a least square regression boost approach that fits regression ensembles to  
 14 minimize mean-squared error. At each step, a new learner is fitted to the difference between the  
 15 observed response and the aggregated prediction of all learners grown previously. The following  
 16 figure represents the algorithm:

17 *Initialize*,  $F_0(x) = \bar{y}$   
 18 *for*  $m = 1$  to  $M$  *do*:  
 19  $\tilde{y}_i = y_i - F_{m-1}(x_i)$ ,  $i = 1, N$   
 20  $(\rho_m, \alpha_m) = \arg \min_{\rho, \alpha} \sum_{i=1}^N [\tilde{y}_i - \rho h(x_i; \alpha)]^2$   
 21  $F_m(x) = F_{m-1}(x) + \rho_m h(x; \alpha_m)$   
 22 *endfor*  
 23 Output the final regression function  $F_m(x)$ .

24 **Fig. 3.** The least square regression boost algorithm (Friedman, 2001).

#### 25 4.3.3. Kalman Filter method

26 ATT is chosen for observation, as it is the observation nearest the DTT to be predicted. After  
 27 rearranging Equation 4 as  $ATT_t = DTT_t - \Delta tt_t$  and treating  $-\Delta tt_t$  as the observation noise  
 28 ( $v_t$ ), observation equation shows a linear relationship between ATT and DTT in Equation 6:

$$29 \quad ATT_t = DTT_t + v_t \quad (6)$$

30 where,  $v$  denotes the observation noise. This equation is equivalent to a standard KF observation  
 31 equation such as  $z_t = Hx_t + v_t$  where  $z$  and  $x$  represents observation and state variables (or  
 32 ATT and DTT in this case) respectively. The observation matrix, H is assumed to be 1.

33 The current traffic condition is more correlated with close conditions than it is with distant  
 34 conditions. The proposed state-space equation is:

$$35 \quad DTT_t = \Phi_{t-1} DTT_{t-1} + w_t \quad (7)$$

36 And the transition function,  $\Phi_{t-1}$ :

$$\Phi_{t-1} = DTT_{t-1} / DTT_{t-2} \quad (8)$$

where  $w$  denotes the state-space noise. The proposed state-space equation is similar to a standard KF state-space model,  $x_t = \Phi_{t-1}x_{t-1} + w_t$  where  $x$ , the state variable, is replaced with the DTT.

The proposed KF model based on (Welch and Bishop, 2006) is described below:

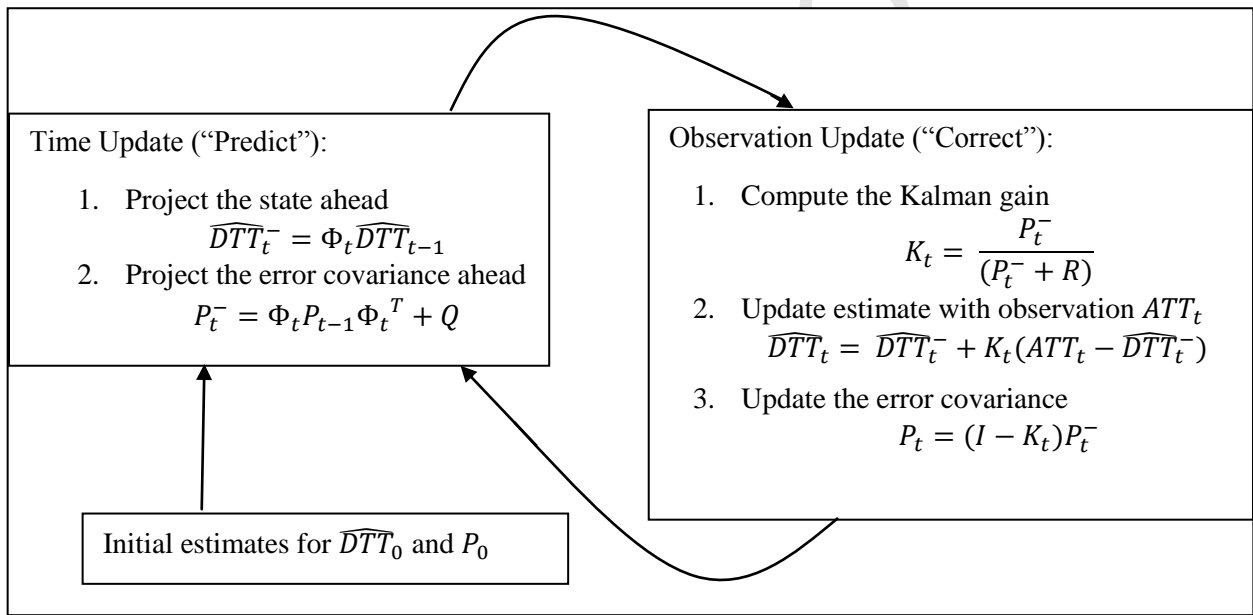


Fig. 4. KF model.

The most recent DTT ( $t - 1$  and  $t - 2$  steps) is unavailable until the vehicles have finished travelling the route; therefore, the DTT from the latest and same historical day of week at  $t - 1$  and  $t - 2$  steps are used to estimate  $\Phi_t$ .  $w$  and  $v$  are assumed to be independent of each other and follow the normal probability distributions:  $p(w) \sim N(0, Q)$  and  $p(v) \sim N(0, R)$ .

A priori ( $\widehat{DTT}_t^-$ ), according to its definition, should be equal to a corresponding DTT. Therefore, the differences between a priori (i.e. state-space projection) and DTT are considered the state-space noise. the State-space noise ( $w_t$ ) is measured by:

$$w_t = DTT_t - \widehat{DTT}_t^- \quad (9)$$

where  $\widehat{DTT}_t^-$  = Corresponding a priori of DTT at time  $t$ .

The observation - after the noise is removed - should be equal to the predicted DTT or a posterior, according to eq. (6). Therefore, the difference between observation and predicted time (i.e. DTT-ATT) is considered as the observation noise. The observation noise ( $v_t$ ) at time  $t$  can be expressed as:

$$v_t = DTT_t - ATT_t \quad (10)$$

Considering computational simplicity and the availability of sufficient data, noise from historic all-days (instead of same-days) was used to estimate the noise covariance. Assumptions regarding the temporal characteristics of noise can be categorized into three types:

- 1 a) Steady Noise (SN): Regardless of time of day, noise is assumed to be the same for a day and  
 2 estimated from the complete training dataset.
- 3 b) Contextual Noise (CN): Noise is assumed to vary by traffic state (free-flow, delay, recurrent,  
 4 and non-recurrent congestions). **Therefore, the entire dataset is divided into four subsets –**  
 5 **free-flow, delay, recurrent, and non-recurrent congestions – based on the travel time. Four**  
 6 **covariance matrices ( $Q_{ff}$ ,  $Q_{dl}$ ,  $Q_{rc}$  and  $Q_{nrc}$ ) are estimated using these sub-datasets.**
- 7 c) Time-varying Noise (TVN): Noise is assumed to vary with every time step of prediction;  
 8 hence, the covariance matrix ( $Q_t$ ) is estimated by the noise of a training dataset at time  $t$ . **In**  
 9 **other words, covariance matrices are estimated by splitting the entire dataset into 1,440**  
 10 **subsets for 1,440 intervals of a day. Therefore,  $Q = \{Q_1, Q_2, \dots, Q_{1440}\}$  depending on  $t^{\text{th}}$**   
 11 **interval of a day.**

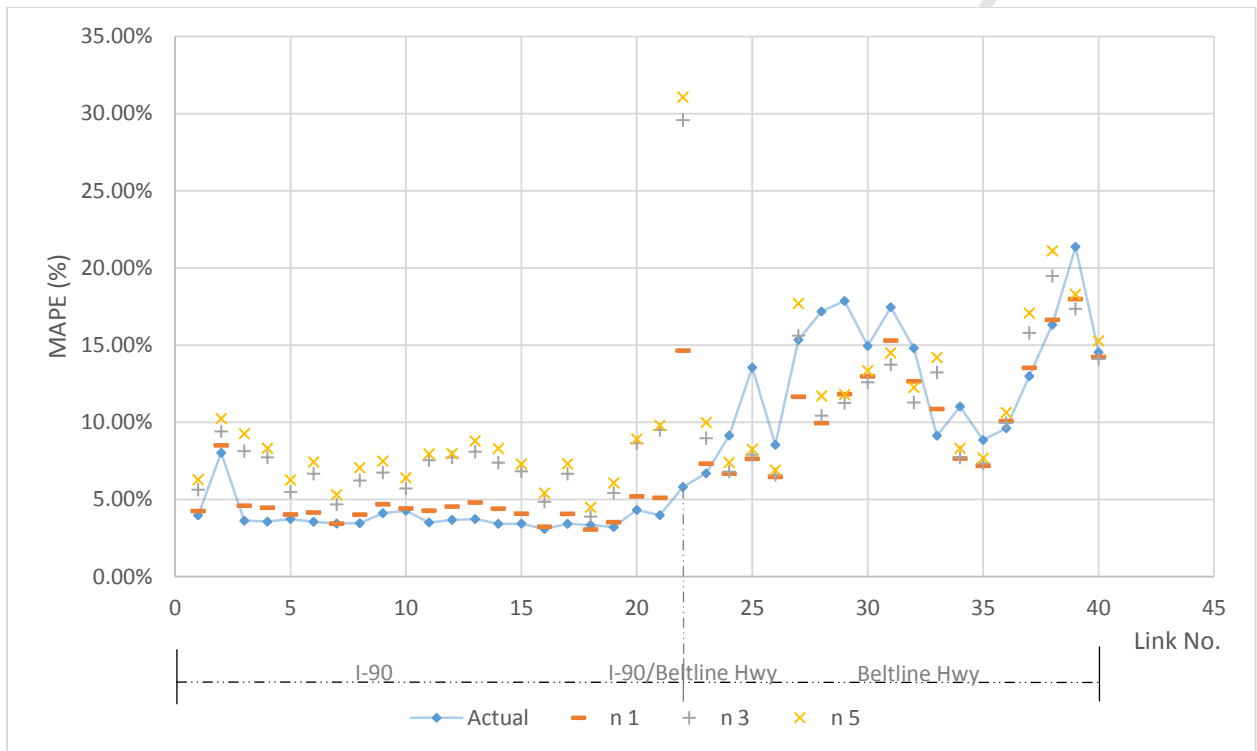
#### 13 4.4. Prediction Performance Evaluation Criteria

14 The mean absolute error (MAE), mean absolute percentage error (MAPE) and root-mean-square  
 15 error (RMSE) were applied to evaluate the performance of travel time prediction methods. In  
 16 general, lower value of the MAE, MAPE and RMSE indicates the superiority in prediction.  
 17 However, a lower MAE, MAPE and RMSE for a corridor that experiences free-flow condition,  
 18 not necessarily prove that the prediction method is superior. Since travel time of such a corridor  
 19 mostly remains unchanged for a short time interval, a naïve method (e.g. using current travel  
 20 time as the 5-min ahead prediction), as opposed to an advanced method, might work best.  
 21 Therefore, it is important to quantify the performance of the naïve method and compare the  
 22 results with the results of advance methods for better understanding. The comparison helps to  
 23 quantify the overall improvements made by an advance method. In this study, the MAE, MAPE  
 24 and RMSE from the prediction results of a naïve method, considering that using ATT as the  
 25 prediction of DTT is a naïve method, was termed as Actual MAE, Actual MAPE and Actual  
 26 RMSE (i.e. AMAE, AMAPE and ARMSE) respectively.

## 28 5. Results & discussion

29 The complete dataset in this study was divided into two sets: Training and Validation.  
 30 Twenty-eight of forty-seven days' worth of data was used for the training dataset, and the rest of  
 31 the data was used as the validation dataset. It was more appropriate to use link speed data as  
 32 opposed to link travel time due to the variability in link lengths. The prediction performance  
 33 index was utilized to perceive global performance (i.e. the performance of the entire network).  
 34 The MAPE was calculated based on the travel time dataset in order to discern the local  
 35 performance at each link. Quantifying improvement is impossible without knowing the AMAPE  
 36 or actual lag/gap, considering that using ATT as the prediction of DTT is a naïve method.  
 37 AMAPE (i.e. MAPE of ATT) was used as the benchmark for MAPEs generated by other  
 38 methods.

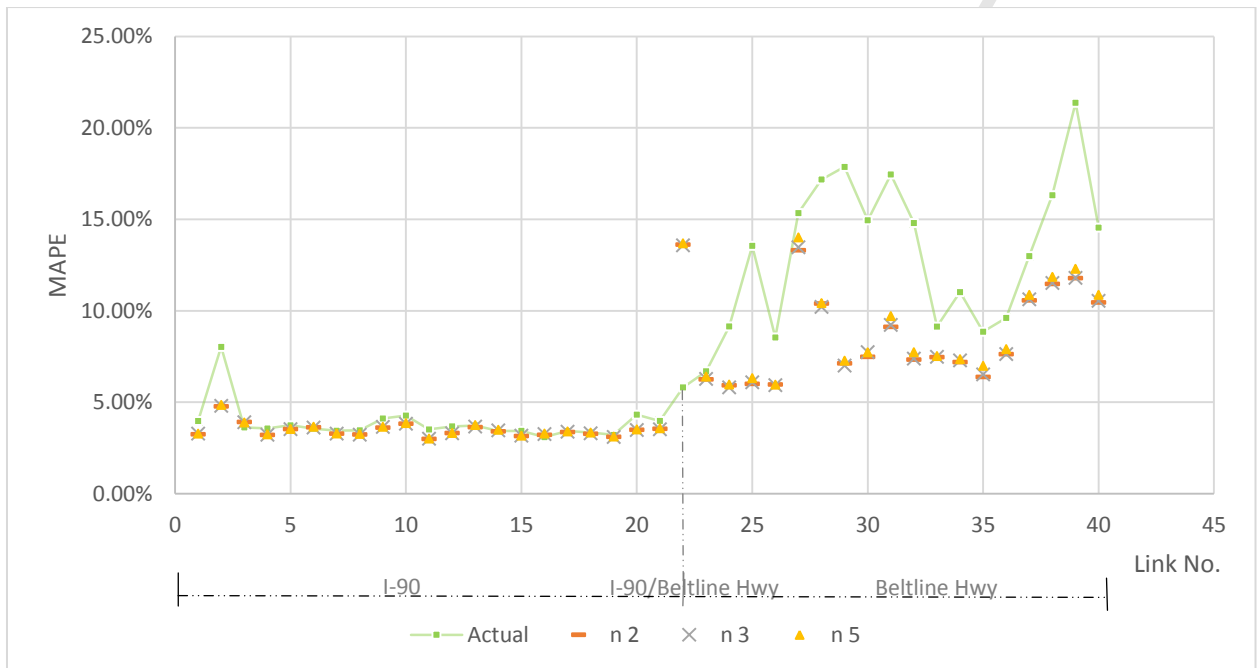
39 The k-NN method showed different prediction performance of different values of  $n$  ( $n = 1, 2,$   
 40  $\dots, 10$ ). [Fig. 5](#) represents the local (i.e. each link) performance of the k-NN model for validation  
 41 dataset when  $n=1, 3$  and  $5$ . When evaluating criteria of local performance, the MAPE of the  
 42 prediction for each link should be smaller than that of the actual lag/gap (AMAPE or the  
 43 benchmark).



**Fig. 5.** MAPE of k-NN model at each link vs. actual gap.

In Fig. 5, MAPEs of travel times predicted by the k-NN method using  $n = 1, 3$  and  $5$  are compared to the actual lag/gap. The local performance is unacceptable since the MAPEs of more than half of the links are greater than the actual lag/gap. Performance shows a decreasing trend (i.e. MAPE increases) with the increase of  $n$ . Moreover, the global performance measured by the MAPE over entire route is poor. The overall MAPEs are higher than the actual gap or AMAPE (6.70%), suggesting the lack of repeating traffic condition over the entire period. Therefore, k-NN is not an appropriate method to predict DTT from ATT when the traffic conditions in the training dataset are not similar to the conditions in the validation dataset.

Two new variables, day of the week and time of the day, were included besides ATT in the LSBoost application. In addition, the latest ATTs (2, 3, 5, 10 minutes etc.) were combined with the current ATT to create a new variable. Similar prediction performances were observed for different numbers of newly added ATT columns ( $n = 2, 3, 5$  and  $10$ ). In other words, this algorithm shows limited sensitivity to the size of the combined ATTs. For visual clarity, only two selected results with the actual gap are shown in Fig. 6 below:

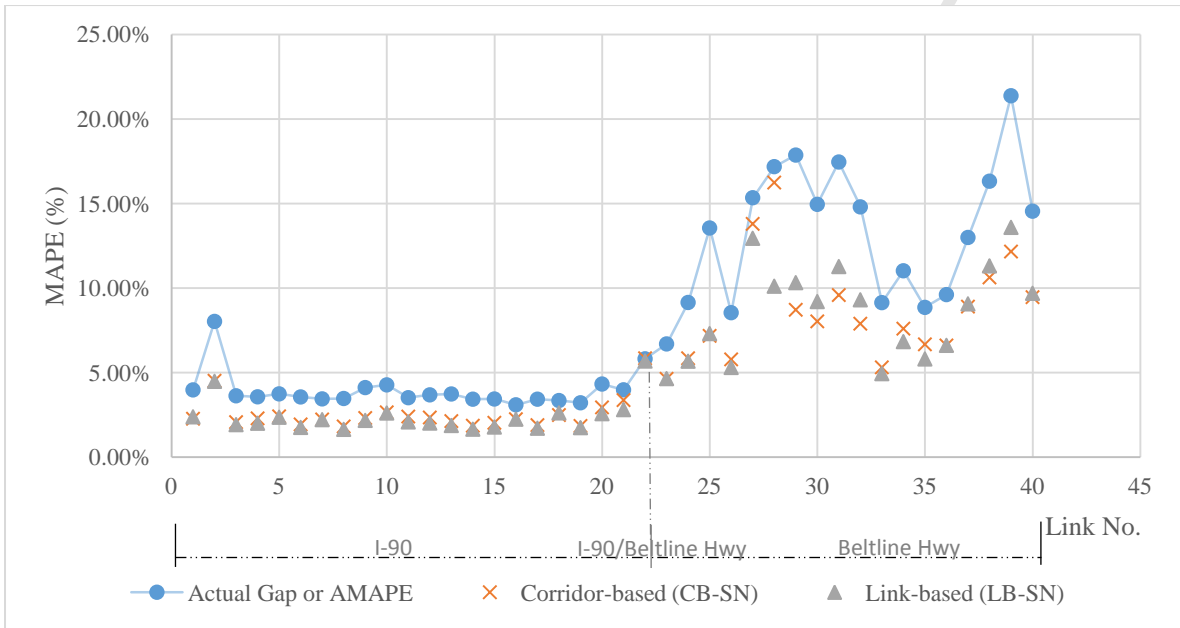


**Fig. 6.** MAPE of LSBoost at each link vs. actual gap.

In Fig. 6, MAPEs of travel times predicted by LSBoost using two, three and five previous-step ATTs ( $n = 2, 3$  and  $5$  respectively) with current ATT are compared to the actual lag/gap. The MAPEs and AMAPEs of different links on I-90 have very similar values which indicate the shortcoming of this method on a freeway that experiences free-flow. The performance of LSBoost prediction is notably well for the Beltline Highway that experiences congestion. The link (22<sup>nd</sup>) that connects I-90 and Beltline corridor performs very poorly showing MAPE much higher than that of AMAPE. Therefore, LSBoost is not recommended at this point to predict DTT from ATT for freeways.

Spatial noise characteristics for KF models were assumed in this study by considering that a) noise of different links in a particular corridor can have similar characteristics, and b) noise of different links, regardless of corridor, can have different characteristics. Previously, three categorical assumptions regarding temporal characteristics of noise have been discussed; therefore, the output of the KF model would be affected by six different estimation procedures of noise covariance regarding the spatial-temporal characteristics of noise. The six methods are: corridor-based steady noise (CB-SN), contextual noise (CB-CN), time-varying noise (CB-TVN), link-based steady noise (LB-SN), contextual noise (LB-CN), and time-varying noise (LB-TVN).

Appropriate noise characteristics of KF should be determined through the evaluation of local performance. Fig. 7, Fig. 8, and Fig. 9 represent the local (i.e. each link) performance of the KF model for validation dataset with different noise assumptions. When evaluating criteria of local performance, the MAPE of the prediction for each link should be smaller than that of the actual lag/gap (AMAPE or the benchmark). In Fig. 7, MAPEs of prediction by the KF model using CB-SN and LB-SN are compared to the actual lag/gap. The prediction performances of the KF model using both CB-SN and LB-SN for each individual link are acceptable, as no link shows a MAPE greater than the actual lag/gap.



1

2

**Fig. 7.** MAPE of KF at each link for steady noise assumption vs. actual gap.

3

4

5

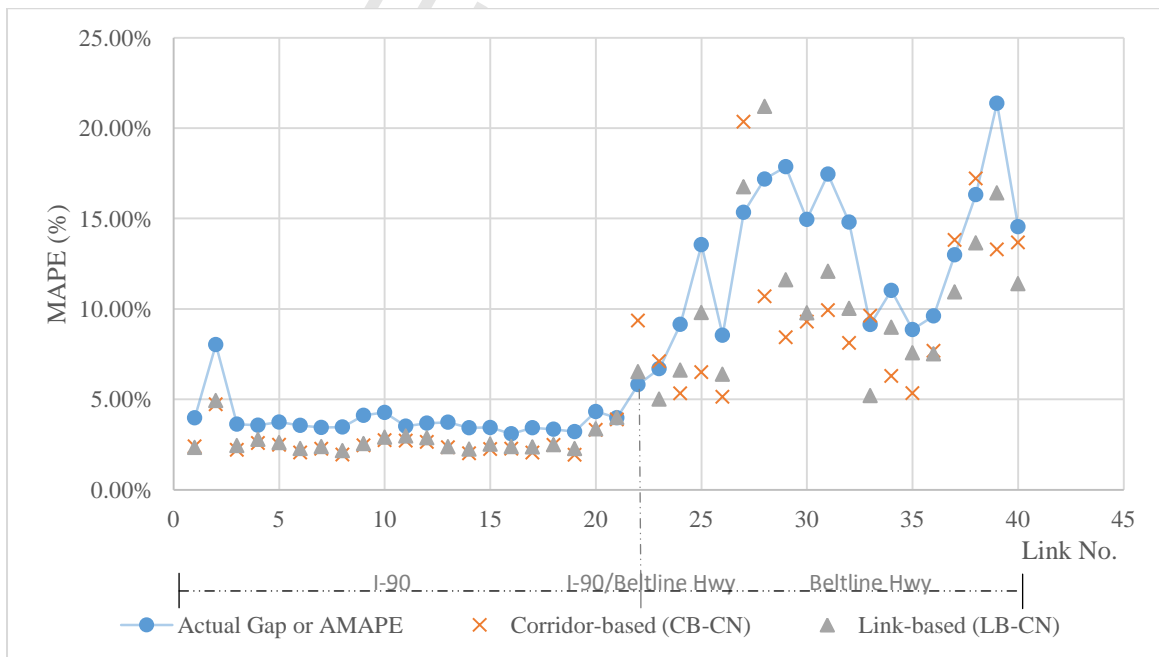
6

7

8

9

In Fig. 8, MAPEs of travel times predicted by the KF model using CB-CN and LB-CN are compared to the actual lag/gap. It is clear that the KF model with both CB-CN and LB-CN has a few links' MAPE greater than AMAPE. In general, the context-based noise assumption is supposed to perform better for the corridor that experiences congestion. The poor performance could be due to the stability of travel time resulting from the saturated traffic flow rate under congestions. At a saturated flow, the variations between ATT and DTT become similar to the variations in free flow or delay conditions.

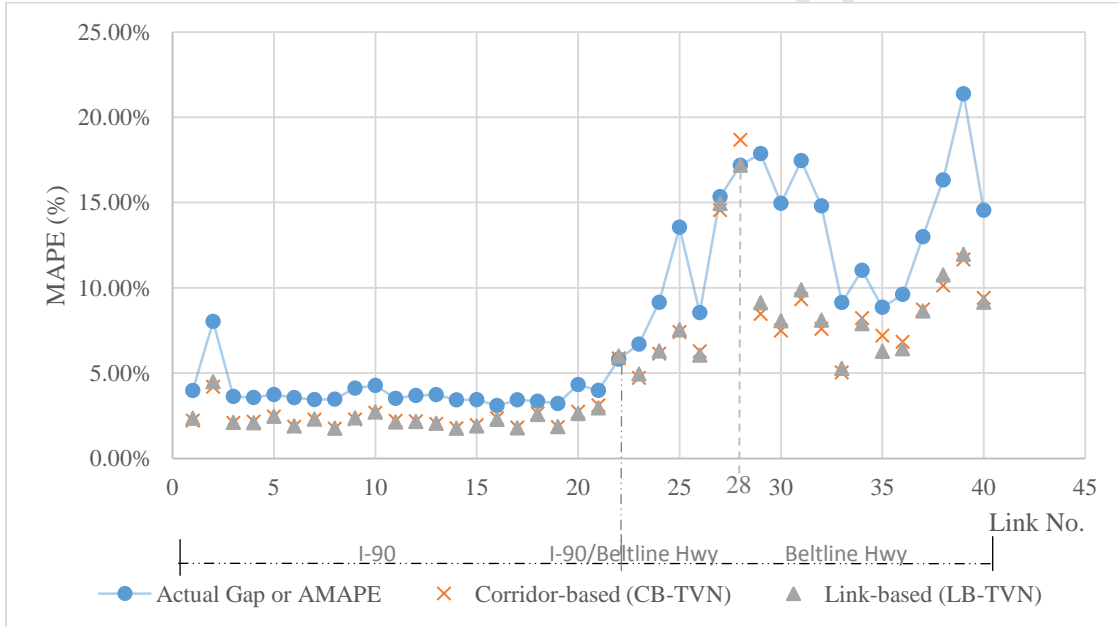


10

11

**Fig. 8.** MAPE of KF at each link for contextual noise assumption vs. actual gap.

1 In Fig. 9, MAPEs of travel times predicted by the KF model using CB-TVN and LB-TVN  
 2 are compared to the actual lag/gap. The assumption of CB-TVN is invalid for a corridor with  
 3 some links that experience traffic congestion. Despite the improvement from contextual noise  
 4 assumption, Fig. 9 shows that the noise homogeneity assumption of CB-TVN is violated at Link  
 5 28, as its performance exceeds the actual lag/gap (AMAPE).



6

7

**Fig. 9.** MAPE of KF at each link for time varying noise assumption vs. actual gap.

8

9 The above discussion provides a comprehensive description of local/link performance of  
 10 different noise assumptions in a KF model. The unambiguous analyses reflect the suitability of  
 11 CB-SN, LB-SN and LB-TVN. Table 1 shows the global performance of prediction expressed by  
 12 the MAE and RMSE of speed data calculated from travel time, which reaffirms the most  
 13 appropriate noise assumption for the dataset is LB-SN.

**Table 1**

14 Overall (global) performance of KF model for selected noise assumptions.

Noise Assumption	Training Dataset		Validation Dataset	
	MAE	RMSE	MAE	RMSE
CB-SN	2.25	6.31	2.41	6.75
LB-SN	2.17	5.96	2.31	6.33
LB-TVN	2.15	6.78	2.45	7.47

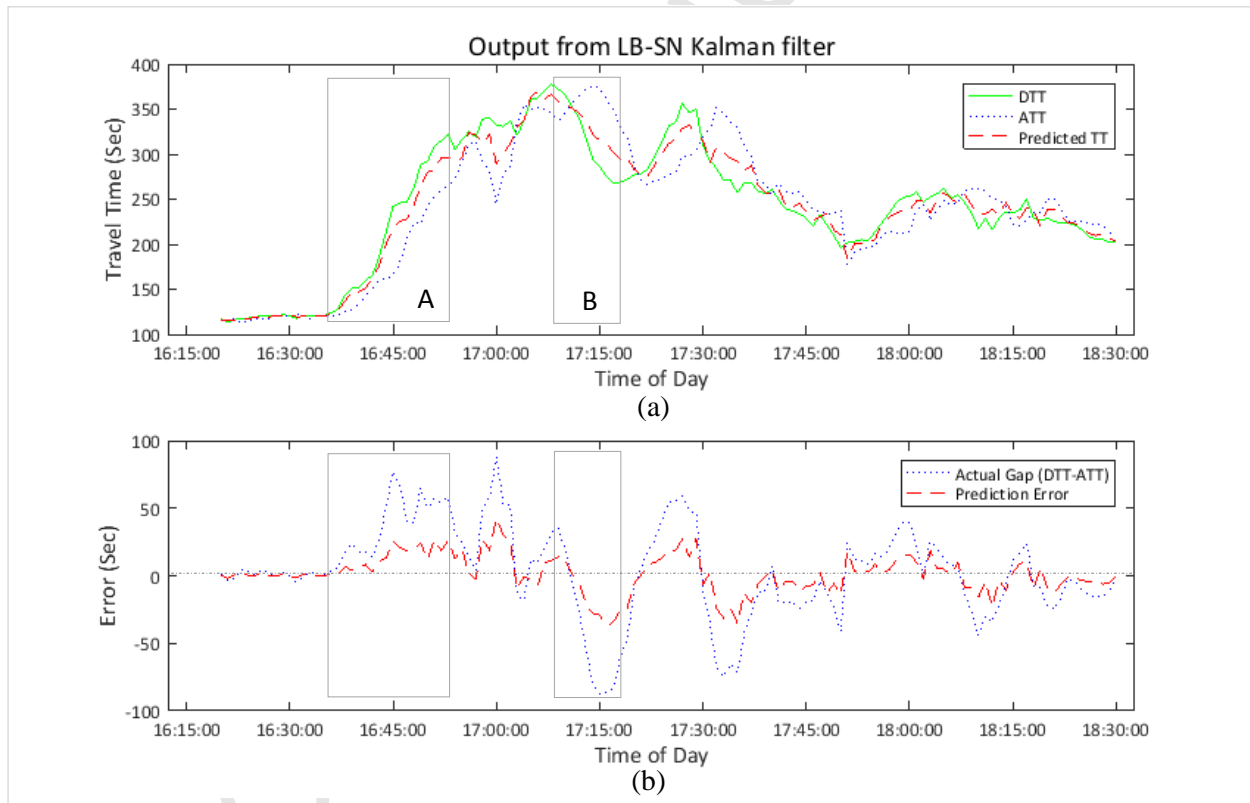
15

16 The KF model with LB-SN assumption is more accurate and computational efficient. For  
 17 instance, the run time of KF with corridor-based noise assumptions is approximately 15mins,  
 18 whereas link-based assumptions take only a minute or two. Corridor-based noise homogeneity  
 19 assumption calls for extra processing of data since Bluetooth data is collected over each link; this  
 20 way, noise covariance can be estimated over the entire corridor. This extra processing increases

1 the run time significantly. The link-based model is the most suitable for an on-line application.  
 2 KF with LB-SN was selected to predict DTT from ATT.

3 The MAPEs of the training and validation datasets using LB-SN KF are 4.27% and 4.53%,  
 4 respectively, whereas, the actual lags/gaps (i.e. AMAPEs) are 6.43% and 6.70%, respectively,  
 5 based on travel time dataset. Improvements are significant when the smaller actual lag/gap is  
 6 considered, including a 50% reduction in AMAPE in some links. When the KF model with LB-  
 7 SN assumption is applied, the validation dataset shows that a 40-50% gap between ATT and  
 8 DTT is minimized in different links of the I-90 corridor, and a nearly 30-40% gap is minimized  
 9 in different links of the Beltline Highway corridor.

10 Moreover, the KF model with the LB-SN assumption has a superior prediction performance  
 11 in cases of traffic state transition (e.g. onset and end of congestion). Fig. 10(a) represents the  
 12 DTT, ATT, and predicted travel time, and Fig. 10(b) demonstrates the actual lag/gap and  
 13 prediction error corresponding to Fig. 10(a).



14

15

**Fig. 10.** Prediction performance between free flow and congested conditions.

16 Fig. 10 clearly depicts that the ATT, DTT, and predicted travel time are almost equal at free  
 17 flow conditions (before 16:30:00). Actual and prediction errors are negligible after 18:15:00  
 18 when congestion is stable. Fig. 10(a) Box A shows the onset of congestion where the actual  
 19 lag/gap is more than 50s, and Fig. 10(b) shows prediction lag is less than 20s. Box B in Fig.  
 20 10(a) shows the end of the congestion situation where the actual and prediction lags show little  
 21 difference (less than 20s). However, at 17:15:00, the mid-point of the end of congestion, the  
 22 ATT is off by 100s from DTT while the prediction error is around 30s. Despite of having a  
 23 moderate improvement (around 40%) according to MAPE (global performance), the prediction



1 shows excellent improvement (around 70%) in cases such as that in box B where there is a state  
 2 transition. Since such cases cover shorter time periods compared to the complete study period,  
 3 the overall performance (local or global) indices are unable to represent the robustness of the  
 4 prediction algorithm. Therefore, the selected noise assumption based on the KF model is capable  
 5 of predicting travel time from the travelers' point of interest (the onset and end of congestion  
 6 rather than free-flow or stable congested conditions).

7 To conclude, to predict the current traffic state as well as travel time, k-NN and LSBoost  
 8 utilize the full dataset whereas the KF method uses most recent dataset. Intrinsically, the  
 9 heteroscedasticity of traffic condition captured by the most recent data is more resemble to the  
 10 current. Perhaps, this gives KF advantages over other methods. Moreover, the noise modeling  
 11 capability offers additional benefits such as the selection of a subset of data to represent the  
 12 outstanding heteroscedasticity that may not be adequately represented by the most recent data  
 13 used in the state projection. Hence, prediction performances of the KF method with noise  
 14 assumptions are more consistent than that of the k-NN and LSBoost in context of different  
 15 roadway links.

## 16

## 17 6. Conclusion

18 ATT (or  $t_{expn}$ ) is the most available form of travel time, but DTT (or  $t_{expt}$ ) is the most  
 19 desirable. The k-NN, LSBoost and KF algorithms were used to predict DTT and thus assist  
 20 motorists by providing a more accurate and reliable travel time. Overall, KF outperformed other  
 21 two methods. Although ATT and DTT differ slightly when the flow of traffic is stable, the  
 22 variation becomes significant when the traffic state is in transition (e.g. moving from unstable to  
 23 stable or vice versa). The KF algorithm with steady noise assumption captured a state of  
 24 transition property accurately and provided an excellent prediction. KF is fast for link-based  
 25 applications, making it desirable for data sources that contain route travel time split into shorter  
 26 links (e.g. loop detectors/Bluetooth data). Although KF is applied (by default) to each link that is  
 27 isolated as a different model, it demonstrates a higher level of accuracy and faster speed due to  
 28 its flexibility, simplicity and compatibility with data characteristics. The application was  
 29 demonstrated during peak periods on freeways covering two corridors – one with fewer  
 30 transitions in traffic state and the other with frequent transitions. Diversity in noise assumptions  
 31 showed negligible impact on the former, while steady noise assumption showed better  
 32 performance on the latter. Steady noise (SN) refers to a fixed covariance estimated from the  
 33 complete training dataset, which indicates that the KF model performs better with the  
 34 generalized noise assumption. Hence, the KF with LB-SN assumption was preferred over other  
 35 assumptions. Better performance was observed during a time when transitions in traffic state  
 36 occurred more frequently.

37 Since arterial highways are supposed to exhibit more state transitions, future research should  
 38 examine the performance of predicting arterial highway travel time to test the robustness of this  
 39 method. Furthermore, as firstly pointed out in Guo et al. 2012, vehicular traffic condition series  
 40 is heteroscedastic in nature, new research is needed to investigate the second-order moment of  
 41 travel time, i.e., the uncertainty of the time series data of travel time. Future research can

1 substantiate and expand on the pioneering research methodologies in the travel time uncertainty  
 2 domain (Guo et al. (2014), Huang et al. (2017), Guo et al. (2017) and Cao et al., (2017)).

### 3 References

- 4 AL-DEEK, H., D'ANGELO, M. P. & WANG, M. Travel time prediction with non-linear time series.  
 5 Fifth International Conference on Applications of Advanced Technologies in Transportation  
 6 Engineering, 1998.
- 7 BACHMANN, C., ABDULHAI, B., ROORDA, M. J. & MOSHIRI, B. 2013. A comparative assessment  
 8 of multi-sensor data fusion techniques for freeway traffic speed estimation using microsimulation  
 9 modeling. *Transportation research part C: emerging technologies*, 26, 33-48.
- 10 BARUTÇUOĞLU, Z. & ALPAYDIN, E. 2003. A comparison of model aggregation methods for  
 11 regression. *Artificial Neural Networks and Neural Information Processing—ICANN/ICONIP*  
 12 2003. Springer.
- 13 BHASKAR, A., CHUNG, E. & DUMONT, A. G. 2011. Fusing loop detector and probe vehicle data to  
 14 estimate travel time statistics on signalized urban networks. *Computer- Aided Civil and*  
 15 *Infrastructure Engineering*, 26, 433-450.
- 16 BLOGG, M., SEMLER, C., HINGORANI, M. & TROUTBECK, R. Travel time and origin-destination  
 17 data collection using Bluetooth MAC address readers. Australasian transport research forum,  
 18 2010.
- 19 CARRION, C. & LEVINSON, D. 2012. Value of travel time reliability: A review of current evidence.  
 20 *Transportation research part A: policy and practice*, 46, 720-741.
- 21 CELIKOGLU, H. B. 2013. Flow-based freeway travel-time estimation: A comparative evaluation within  
 22 dynamic path loading. *IEEE Transactions on Intelligent Transportation Systems*, 14, 772-781.
- 23 CHEN, H. & RAKHA, H. A. 2014. Real-time travel time prediction using particle filtering with a non-  
 24 explicit state-transition model. *Transportation Research Part C: Emerging Technologies*, 43,  
 25 112-126.
- 26 CHEN, M. & CHIEN, S. 2001. Dynamic freeway travel-time prediction with probe vehicle data: Link  
 27 based versus path based. *Transportation Research Record: Journal of the Transportation*  
 28 *Research Board*, 157-161.
- 29 CHIEN, S. I.-J. & KUCHIPUDI, C. M. 2003. Dynamic travel time prediction with real-time and historic  
 30 data. *Journal of transportation engineering*, 129, 608-616.
- 31 CLARK, S., GRANT-MULLER, S. & CHEN, H. 2002. Cleaning of matched license plate data.  
 32 *Transportation Research Record: Journal of the Transportation Research Board*, 1-7.
- 33 COMMANDEUR, J. J. & KOOPMAN, S. J. 2007. *An introduction to state space time series analysis*,  
 34 OUP Oxford.
- 35 DARWISH, D. 2013. Assessment of Offline Digital Signature Recognition Classification Techniques.  
 36 *International Journal of Computer Networks & Communications Security*, 1.
- 37 DION, F. & RAKHA, H. 2006. Estimating dynamic roadway travel times using automatic vehicle  
 38 identification data for low sampling rates. *Transportation Research Part B: Methodological*, 40,  
 39 745-766.
- 40 DURBIN, J. & KOOPMAN, S. J. 2012. *Time series analysis by state space methods*, Oxford University  
 41 Press.
- 42 ELHENAWY, M., EL-SHAWARBY, I. & RAKHA, H. 2017. Modeling the Perception Reaction Time  
 43 and Deceleration Level for Different Surface Conditions Using Machine Learning Techniques.  
 44 *Advances in Applied Digital Human Modeling and Simulation*. Springer.
- 45 FEI, X., LU, C.-C. & LIU, K. 2011. A bayesian dynamic linear model approach for real-time short-term  
 46 freeway travel time prediction. *Transportation Research Part C: Emerging Technologies*, 19,  
 47 1306-1318.
- 48 FRIEDMAN, J. H. 2001. Greedy function approximation: a gradient boosting machine. *Annals of*  
 49 *statistics*, 1189-1232.

- 1 GHOSH, B., ASIF, M. T., DAUWELS, J., CAI, W., GUO, H. & FASTENRATH, U. Predicting the  
2 duration of non-recurring road incidents by cluster-specific models. *Intelligent Transportation*  
3 *Systems (ITSC)*, 2016 IEEE 19th International Conference on, 2016. IEEE, 1522-1527.
- 4 GUO, J., HUANG, W. & WILLIAMS, B. M. 2014. Adaptive Kalman filter approach for stochastic short-  
5 term traffic flow rate prediction and uncertainty quantification. *Transportation Research Part C:*  
6 *Emerging Technologies*, 43, 50-64.
- 7 HAGHANI, A., HAMED, M., SADABADI, K., YOUNG, S. & TARNOFF, P. 2010. Data collection of  
8 freeway travel time ground truth with bluetooth sensors. *Transportation Research Record:*  
9 *Journal of the Transportation Research Board*, 60-68.
- 10 HAMED, M. M., AL-MASAEID, H. R. & SAID, Z. M. B. 1995. Short-term prediction of traffic volume  
11 in urban arterials. *Journal of Transportation Engineering*, 121, 249-254.
- 12 HAMILTON, J. D. 1994. *Time series analysis*, Princeton university press Princeton.
- 13 HARVEY, A. C. 1990. *Forecasting, structural time series models and the Kalman filter*, Cambridge  
14 university press.
- 15 JENELIUS, E. & KOUTSOPOULOS, H. N. 2017. Urban Network Travel Time Prediction Based on a  
16 Probabilistic Principal Component Analysis Model of Probe Data. *IEEE Transactions on*  
17 *Intelligent Transportation Systems*.
- 18 JIANG, W. Is regularization unnecessary for boosting? AISTATS, 2001a.
- 19 JIANG, W. Some theoretical aspects of boosting in the presence of noisy data. Proceedings of the  
20 Eighteenth International Conference on Machine Learning, 2001b. Citeseer.
- 21 JIAO, F., XU, J., YU, L. & SCHUURMANS, D. Protein fold recognition using the gradient boost  
22 algorithm. *Computational Systems Bioinformatics Conference*, 2006. 43-53.
- 23 KHATTAK, A., POLYDOROPOULOU, A. & BEN-AKIVA, M. 1996. Modeling revealed and stated  
24 pretrip travel response to advanced traveler information systems. *Transportation Research*  
25 *Record: Journal of the Transportation Research Board*, 46-54.
- 26 KIM, J., RHO, J. & PARK, D. 2009. On-line estimation of departure time-based link travel times from  
27 spatial detection system. *International Journal of Urban Sciences*, 13, 63-80.
- 28 LIU, H. 2008. *Travel time prediction for urban networks*, TU Delft, Delft University of Technology.
- 29 LIU, H., VAN ZUYLEN, H., VAN LINT, H. & SALOMONS, M. 2006. Predicting urban arterial travel  
30 time with state-space neural networks and Kalman filters. *Transportation Research Record:*  
31 *Journal of the Transportation Research Board*, 99-108.
- 32 MAYBECK, P. S. 1990. The Kalman filter: An introduction to concepts. *Autonomous robot vehicles*.  
33 Springer.
- 34 MOGHADDAM, S. & HELLINGA, B. 2013. Quantifying Measurement Error in Arterial Travel Times  
35 Measured by Bluetooth Detectors. *Transportation Research Record: Journal of the*  
36 *Transportation Research Board*, 111-122.
- 37 MOGHADDAM, S. & HELLINGA, B. 2014. Algorithm for detecting outliers in Bluetooth data in real  
38 time. *Transportation Research Record: Journal of the Transportation Research Board*, 129-139.
- 39 MOSCOSO- LÓPEZ, J. A., TURIAS, I., JIMÉNEZ- COME, M. J., RUIZ- AGUILAR, J. J. &  
40 CERBÁN, M. D. M. 2016. A two- stage forecasting approach for short- term intermodal freight  
41 prediction. *International Transactions in Operational Research*.
- 42 MYUNG, J., KIM, D.-K., KHO, S.-Y. & PARK, C.-H. 2011. Travel time prediction using k nearest  
43 neighbor method with combined data from vehicle detector system and automatic toll collection  
44 system. *Transportation Research Record: Journal of the Transportation Research Board*, 51-59.
- 45 NANTHAWICHIT, C., NAKATSUJI, T. & SUZUKI, H. 2003. Application of probe-vehicle data for  
46 real-time traffic-state estimation and short-term travel-time prediction on a freeway.  
47 *Transportation Research Record: Journal of the Transportation Research Board*, 49-59.
- 48 ODA, T. An algorithm for prediction of travel time using vehicle sensor data. *Road Traffic Control*,  
49 1990., Third International Conference on, 1990. IET, 40-44.
- 50 ODURO, S. D., METIA, S., DUC, H., HONG, G. & HA, Q. 2015. Multivariate adaptive regression  
51 splines models for vehicular emission prediction. *Visualization in Engineering*, 3, 13.

- 1 PUCKETT, D. D. & VICKICH, M. J. 2010. Bluetooth®-based travel time/speed measuring systems  
2 development.
- 3 RICE, J. & VAN ZWET, E. 2004. A simple and effective method for predicting travel times on freeways.  
4 *Intelligent Transportation Systems, IEEE Transactions on*, 5, 200-207.
- 5 SKABARDONIS, A. & GEROLIMINIS, N. 2005. Real-time estimation of travel times on signalized  
6 arterials.
- 7 SMITH, B. L., WILLIAMS, B. M. & OSWALD, R. K. 2002. Comparison of parametric and  
8 nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging*  
9 *Technologies*, 10, 303-321.
- 10 SUMALEE, A., PAN, T., ZHONG, R., UNO, N. & INDRA-PAYOONG, N. 2013. Dynamic stochastic  
11 journey time estimation and reliability analysis using stochastic cell transmission model:  
12 Algorithm and case studies. *Transportation Research Part C: Emerging Technologies*, 35, 263-  
13 285.
- 14 TOPPEN, A. & WUNDERLICH, K. 2003. *Travel time data collection for measurement of advanced*  
15 *traveler information systems accuracy*, Mitretek Systems.
- 16 VAN HINSBERGEN, C., VAN LINT, J. & SANDERS, F. Short term traffic prediction models.  
17 PROCEEDINGS OF THE 14TH WORLD CONGRESS ON INTELLIGENT TRANSPORT  
18 SYSTEMS (ITS), HELD BEIJING, OCTOBER 2007, 2007.
- 19 VLAHOGIANNI, E. I., KARLAFTIS, M. G. & GOLIAS, J. C. 2014. Short-term traffic forecasting:  
20 Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*,  
21 43, 3-19.
- 22 WELCH, G. & BISHOP, G. 2006. An introduction to the kalman filter. Department of Computer Science,  
23 University of North Carolina. Chapel Hill, NC, unpublished manuscript.
- 24 WU, C.-H., HO, J.-M. & LEE, D.-T. 2004. Travel-time prediction with support vector regression.  
25 *Intelligent Transportation Systems, IEEE Transactions on*, 5, 276-281.
- 26 XIAO, Y., QOM, S., HADI, M. & AL-DEEK, H. 2014. Use of Data from Point Detectors and Automatic  
27 Vehicle Identification to Compare Instantaneous and Experienced Travel Times. *Transportation*  
28 *Research Record: Journal of the Transportation Research Board*, 95-104.
- 29 YANG, J.-S. Travel time prediction using the GPS test vehicle and Kalman filtering techniques.  
30 American Control Conference, 2005. Proceedings of the 2005, 2005. IEEE, 2128-2133.
- 31 ZHAN, X., HASAN, S., UKKUSURI, S. V. & KAMGA, C. 2013. Urban link travel time estimation  
32 using large-scale taxi data with partial information. *Transportation Research Part C: Emerging*  
33 *Technologies*, 33, 37-49.
- 34 ZHANG, X., ONIEVA, E., PERALLOS, A., OSABA, E. & LEE, V. C. 2014. Hierarchical fuzzy rule-  
35 based system optimized with genetic algorithms for short term traffic congestion prediction.  
36 *Transportation Research Part C: Emerging Technologies*, 43, 127-142.
- 37 ZHANG, Y., HAGHANI, A. & ZENG, X. 2015. Component GARCH models to account for seasonal  
38 patterns and uncertainties in travel-time prediction. *IEEE Transactions on Intelligent*  
39 *Transportation Systems*, 16, 719-729.
- 40 ZHENG, F. & VAN ZUYLEN, H. 2013. Urban link travel time estimation based on sparse probe vehicle  
41 data. *Transportation Research Part C: Emerging Technologies*, 31, 145-157.
- 42 ZHONG, R., LUO, J., CAI, H., SUMALEE, A., YUAN, F. & CHOW, A. H. 2017. Forecasting journey  
43 time distribution with consideration to abnormal traffic conditions. *Transportation Research Part*  
44 *C: Emerging Technologies*, 85, 292-311.
- 45 ZOU, Y., ZHU, X., ZHANG, Y. & ZENG, X. 2014. A space-time diurnal method for short-term freeway  
46 travel time prediction. *Transportation Research Part C: Emerging Technologies*, 43, 33-49.
- 47 GUO, J., HUANG, W., & WILLIAMS, B.M., 2012, Integrated heteroscedasticity test for vehicular  
48 traffic condition series, ASCE Journal of Transportation Engineering, Vol.138, Issue 9, pp. 1161-  
49 1170.

- 1 GUO, J., LIU, Z., HUANG, W., WEI, Y., & CAO, J., 2017, Short term traffic flow prediction using fuzzy  
2 information granulation approach under different time intervals, IET Intelligent Transport  
3 Systems, in press
- 4 HUANG, W., JIA, W., GUO, J., WILLIAMS, B.M., SHI, G., WEI, Y., & CAO, J., 2017, Real time  
5 prediction of seasonal heteroscedasticity in vehicular traffic flow series, IEEE Transactions on  
6 Intelligent Transportation Systems, in press
- 7 CAO, J., LI, R., HUANG, W., GUO, J., & WEI, Y., 2017, Traffic network equilibrium problems with  
8 demands uncertainty and capacity constraints of arcs by scalarization approaches, SCIENCE  
9 CHINA Technologica Sciences, accepted, doi:10.1007/s11431-017-9172-4  
10

ACCEPTED MANUSCRIPT