Applied Soft Computing 72 (2018) 65–78

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions

Soknath Mil, Mongkut Piantanakulchai*

School of Civil Engineering and Technology, Sirindhorn International Institute of Technology, Thammasat University, P.O. Box 22, Pathum Thani, 12121, Thailand

ARTICLE INFO

Article history:
Received 15 November 2017
Received in revised form 18 June 2018
Accepted 25 June 2018
Available online 17 July 2018

Keywords:
Bayesian data fusion approach
Gaussian mixture model
Travel time estimation

ABSTRACT

This paper presents a framework for the development of the travel time estimation model using multiple sources of data with consideration of spurious data and traffic conditions. A modified Bayesian data fusion approach, combined with the Gaussian mixture model, is used to fuse the travel time data, which are estimated from different types of sensors to improve accuracy, precision, as well as completeness of data, in terms of spatial and temporal distribution. Two additional features are added into existing models including the difference of traffic conditions classified by the Gaussian mixture model and the bias estimation from individual sensor by introducing a non-zero mean Gaussian distribution which learned from the training dataset. The methodology and computational procedure are presented. The Gaussian mixture model is used to classify states of traffic into predefined number of traffic regimes. Once a traffic condition is classified, the modified Bayesian data fusion approach is used to estimate travel time. The proposed model provides explicit advantages over the basic Bayesian approach, such as being robust to noisy data, reducing biases of an individual estimation, and producing a more precise estimation of travel time. Two different real-world datasets and one simulated dataset are used to evaluate the performance of the proposed model under three different traffic regimes: free flow, transitional flow and congested flow regimes. The results when compared with the results from benchmark models show significant improvement in the accuracy of travel time estimation in terms of mean absolute percentage errors (MAPE) in the range of 3.46% to 16.3%.

1. Introduction

Recently, traffic congestion has become a major concern for most urbanized areas in the world. It generates various adverse effects. Perhaps one of the most important issues is its impediment to economic growth as it restrains smooth operation of transportation activities. Common practices to alleviate traffic congestion are to apply traffic policies, such as expansion of road capacity, encourage the use of public transport and use congestion pricing. Travel time is an indicator of traffic conditions and system performance for road authorities [1], as well as information which is easy to understand by road users [2]. However the benefits of reducing uncertainty of travel time information are often disregarded [3]. Uncertainty of travel time information implies a higher uncertainty; the additional time road users need to reserve for their journey to avoid being late. Reducing uncertainty means to improve reliability or in other words, making travel time information more precise. Reduction of travel time uncertainty is a key of the successful transportation system management, especially for an Advance Traveler Information System (ATIS). The ATIS can significantly improve the reliability of travel time estimation by collect traffic data, using various modern technologies (Wi-Fi, Bluetooth, CCTV, etc.) with an appropriate data fusion model. However, several important matters have to be discussed before any fusion model is selected.

In general, travel time estimation based on a single type of data suffers from two main factors: *the quality of traffic data,* and *the travel time estimation model itself*. The quality of traffic data depends on many factors, such as spacing of fixed sensors (loop detectors, CCTVs, etc.), penetration rate of moving sensors (mobile phones, GPS, etc.), intrinsic error of sensors (mechanical deformation due to temperature, and positioning errors of GPS device, etc.). There are also a malfunctions of communication and controller devices [4]. Another possibility of errors is from a model misspecification, which varies among different selected models according to the nature of the data. For example, conventional methods of speed

* Corresponding author.
E-mail address: mongkut@siit.tu.ac.th (M. Piantanakulchai).

https://doi.org/10.1016/j.asoc.2018.06.046
1568-4946/

based travel time estimation tend to underestimate travel time during congested flow [5].

Nowadays, even though technology development helps traffic engineers to collect data through various types of sensors, most of the existing infrastructures use point detections of traffic (fixed sensors), especially inductive loop detectors [6]. The spacing between loop detectors may range from 0.8 km [7] to 3.5 km [8] due to their high cost of installation. Generally, most inductive loop detectors are single loop detectors, only capable of counting vehicles and measuring occupancy. According to this limitation in measuring speed, the common techniques found in the literature use the assumption of an average vehicle length to derive a fixed speed over the link, and then, the speed based travel time estimation model is applied. Moreover, advanced models based on traffic flow theory allow the relationship among traffic variables to be derived, which lead to the calculation of the travel time. Generally, these methods of estimation are based on flow conservation and propagation principles [9]. Additionally, statistical learning approaches (data-based models) applied on single loop detector data, are also found in the literature. Various techniques such as an artificial neural network, polynomial regression model, time series model, and stochastic regression model, are used to get the relationship among variables (flow and density as related to travel time, in this case). Unlike single loop detectors, dual loop detectors are able to capture more traffic variables, including volume, presence, occupancy, speed, headway, and gaps. With these parameters, various travel time estimation methods are applied, such as the trajectory reconstruction method, vehicle re-identification method, traffic theory based method, and data-based method [10]. However, both types of fixed sensors (single and double loop detectors) have a common problem, the inability to capture area-wide dynamic behavior of traffic, which leads to the problem of either overestimation or underestimation of the travel time [4]. In summary, the travel time estimation methods based on inductive loop data suffer from the wide spacing between sensors and the assumption used in the calculation, i.e. the assumptions of vehicle length in conventional speed based models, and the assumption of gap distance in the traffic flow models.

On the other hand, GPS and cell phone device usage are sharply increasing nowadays, with the ability to collect speeds and locations at different timestamps. This provides alternative data for the travel time estimation model, as seen in the literature. In general, travel time estimated from GPS or cell phone equipped vehicles (probe vehicles) is calculated by averaging the travel time data recorded from a number of probe vehicles. However, it is obvious that the accuracy of travel time estimated from this method is biased since the availability of the probe vehicle data is small and does not represent the whole population. A review of more complex statistical techniques can be found in [10]. These include a combination of current travel time and historical travel time as a linear combination, and the use of the Bayesian conjugate and regression models. Beside statistical models, other methods, including fuzzy logic, Markov chains, and dynamic Bayesian network, are also found in the literature. Even though these various techniques attempted to improve the accuracy of travel time estimation, the accuracy of these estimation methods is low when the amount of probe vehicles (penetration rate) is limited in the traffic stream, as stated in [11,12]. Other factors, which may influence the accuracy of estimation, include loss of signal, positioning errors, and driving behavior.

In short, travel time estimation methods described above are flawed differently according to data from different types of sensors and the calculation procedures. Thus, there is strong motivation for developing a data fusion model in which these issues are taken into consideration. Most of the previous studies, according to the review by [10], use only one type of data as a model input. In spite of that, the objective of this paper focuses on the improvement of travel time estimation using the data fusion model. The following literature review discusses those applied data fusion models.

2. Related works

Interest in data fusion research begun in the late 90 s [13]. Various fusion methodologies and architectural frameworks have been proposed. One of the first data fusion models proposed by [14] under the Advanced Driver and Vehicle Advisory Navigation Concept (ADVANCE) project, relied on a simple convex combination of the observations. Similar studies used the Bar-Shalom/Compo combination, in which the covariance between estimators is taken into account [15,16]. Generally, these combination techniques apply weights to estimators (sensors) according to an individual statistical property (e.g. sum squares of estimated errors). One of the disadvantages of this method is that fixed weights are assigned throughout the estimation which is not flexible.

Another technique is the use of the Kalman filter, a recursive process of estimation. This technique requires two fundamental equations to represent the process and the measurement, called state-space equations. Different state-space equations were modified to incorporate different types of traffic data in [4,16–19]. Usually, traffic data are provided in different frequency or even missed, which are the limitations of the basic Kalman filter model. A modified version of the Kalman filter, called SCAAT, was proposed by [20]) in order to relax these limitations. The performance of the Kalman filter basically depends on how the process is modelled. A more complex process model tends to improve the results of travel time estimation, compared to the simple random walk model [16].

In addition to the Kalman filter, a number of neural network based approaches can be found in [21–24]. They consist of simple elements (nodes), processing in a parallel manner. Nodes are connected by links represented by different weights, in which the values are obtained by training. Different numbers of hidden layers are modelled, depending on the complexity of the problem to be solved. One of the advantages of a neural network is its ability to handle non-linear relationship among inputs. In the application of traffic research, several feed forward neural networks are presented by considering the input nodes from different sensors, for example, loop detector data and probe vehicle data. However, the calculation process of a neural network is relatively a black-box. Thus, the importance of inputs is not easily assessed, and users may fail to identify unimportant predictor variables [25].

Finally, Bayesian theory is another interesting technique which has been applied to the data fusion model. Application of Bayesian theory was applied to the study of [26]. In their study, two types of traffic data were used, loop detectors and toll ticket data. A similar approach that made use of probabilistic theory, Evidential Theory or Dempster-Shafer was proposed by [27,28,33]. Evidential theory is the generalization of Bayesian theory in which the possible discrete states and the combination of states is considered by credibility as a belief function.

From the literature review on travel time estimation using the data fusion models above suggests that even though compromising results are achieved in the related studies, those works have their own limitations because of their assumptions and modelling frameworks. Without appropriate de-noising methods, travel time estimation can be severely biased. Thus, the motivations of this study to develop an efficient data fusion model for travel time estimation are as follows: a) Travel time estimation from only one type of sensor suffers from various issues. These issues create different reliability of estimations during different traffic regimes. b) There is a number of studies conducted to improve travel time estimation. However, few studies use data fusion techniques. c) Most of the

data fusion based travel time estimation models are invariant to different traffic regimes. d) Most of the models do not consider different reliability levels of data at the point of estimation. e) Finally, most of the models only provide a point estimated value without its level of reliability.

The concept of the Bayesian approach in data fusion, applied to traffic data is not new, presented by [26]. However, in his study, only travel time from two types of sensors are considered, and the variation of errors at different time intervals is not taken into account in the formulation of the sensor model.

In this study, the Bayesian approach is adapted to handle the issues found in the literature. The proposed fusion model, based on Bayesian theory, uses the advantages of by each type of data source, attempting to produce more reliable travel time estimation, overcoming individual limitations of data sources. Two distinct benefits, based on the concept of the Bayesian approach, applied to travel time estimations, as compared to other approaches, can be seen clearly. Firstly, estimated travel time using the Bayesian approach can improve the accuracy and precision of estimation, in terms of mean and variance. Secondly, unlike some models presented in the literature, the use of the Bayesian approach is not limited to a specific type of flow. For example, the traffic flow model may vary between interrupted and uninterrupted flow.

In this study, we acknowledge the nature of travel time data which always contain errors, especially during congested conditions. Motivated by [29], we proposed a methodology based on the Bayesian data fusion model by incorporating two additional features into existing model including the difference of traffic conditions classified by the Gaussian mixture (GM) model (Section 5) and the bias estimation from individual sensor by introducing a non-zero mean Gaussian distribution (Section 4.1) which is learned from the training dataset. Regarding the first feature, different conditions of traffic information were taken into account so that the proposed model would be supplied with appropriate parameters at specific traffic condition. With respect to the second feature, the "Non-zero means" represent noises with biases of sensors (Eqs. 4 and 5) which might occur from various sources of errors (i.e. long distance of detector placement would produce underestimation of traffic condition if an accident happens in the middle segment or insufficient GPS data penetration rate to statistically represent traffic condition). A combination of the Bayesian approach and GM model has advantages over the pure Bayesian model, since it is robust to the noise which is generated differently during different traffic conditions. The proposed model produces the dynamic weights of sensors according to the traffic conditions and outliers of the estimations.

The rest of the paper is managed as follows. Firstly, the approach of the Bayesian data fusion proposed by [29], is presented. In this study, the approach is adapted to combine travel time from different sensors for this study. Secondly, the Gaussian mixture model as a traffic classifier is presented. Three different case studies are investigated to examine the models with different traffic assumptions. One of the case studies used the simulated data under the AIMSUN program environment while the other two case studies applied real-world data. One set of data made use of the raw data collected by the Mobile Century Project [30] while the other set of data uses the link travel time presented in the study by [31]. Finally, the results and discussion of individual model during different traffic conditions are discussed.

3. Modelling framework

One of the most important concepts in this study is that there exist different error distributions among different traffic conditions. Thus, the de-noising of data is made in different traffic conditions, in terms of additive Gaussian noise, to the estimation. Mean and variance of estimated errors during different traffic conditions is a key process in deriving the correct model parameters. Fig. 1 describes the online operation scheme of the proposed methodology. At each time step, the evaluation of a traffic condition is performed by using the posterior probability of the Gaussian mixture model in which its parameters are defined by the Expectation-Maximization (EM) algorithm during the training stage. The training outputs are the travel time distribution model parameters (GM model parameters) and its corresponding parameters of an individual sensor (means and variances). These parameters are used to apply in the Bayesian data fusion model to produce the posterior probability. Finally, the posterior probability is maximized in which the outputs are obtained, in terms of the estimated travel time (mean) and the variation of estimation (variance). Detailed procedures are explained in the following sub-sections.

Pseudo code representing the above process can be written as:

```
While true:
    # get traffic data from N sensors
    Get S = [S1[t], S2[t], … Sn[t]]
    # get Medium value among traffic sensor data
    Set Medium = medium(S)
    # Identified traffic state using trained Gaussian Mixture Model
    Set State = GMM(Medium)
    # apply Bayesian Fusion model with respect to traffic
    # State, selected model (Spurious and Non-Spurious
    # Model), bias and standard deviation of estimation from
    # individual sensor (received from training stage) to Estimated
    # Travel Time (ETT)
    ETT = BF (S, Bias(S), Standard(S), State, Model);
    Return ETT
    # Wait for next incoming sensor data
    Sleep ()
```

Step 1: Training stage: travel time data are separated into two sets: training and validating. The training set is used to fit the Gaussian mixture model according to a predefined set of traffic conditions. These parameters are then used to determine whether the travel time is in which traffic condition (during the training stage and estimation stage). Thus, the training data is finally grouped into different traffic conditions, and the error distribution of estimation is found accordingly. Lastly, the outputs from the training stage are the GM model parameters, and the error distribution parameters (means, variances) for individual sensor types of different traffic conditions.

Step 2: Estimation stage — at each interval (i.e. 1 min, 5 min), when the travel time that are estimated from different type of sensors are available, their median value is used to indicate the degree of each traffic condition according to the posterior probability provided by the GM model. The median is used because it outperforms the average without knowing the statistical characteristics of data [32]. Then, the modified Bayesian data fusion approach is used, with respect to the parameters defined during the training stage.

4. Bayesian approach in data fusion model (Box 1)

Bayesian inference is a statistical inference technique, relying on Bayes' theorem. In the Bayesian data fusion approach, multiple observations are used to infer the true state of the object that we are estimating using Bayes' theorem. It is the probability of an event based on the conditions related to the event, which is stated in the following equation:

$$P(X = x \mid S = s) = \frac{P(S = s \mid X = x) \cdot P(X = x)}{\int P(S = s \mid X = x) \cdot P(X = x)\, dx} = \frac{P(S = s \mid X = x) \cdot P(X = x)}{P(S = s)} \quad (1)$$

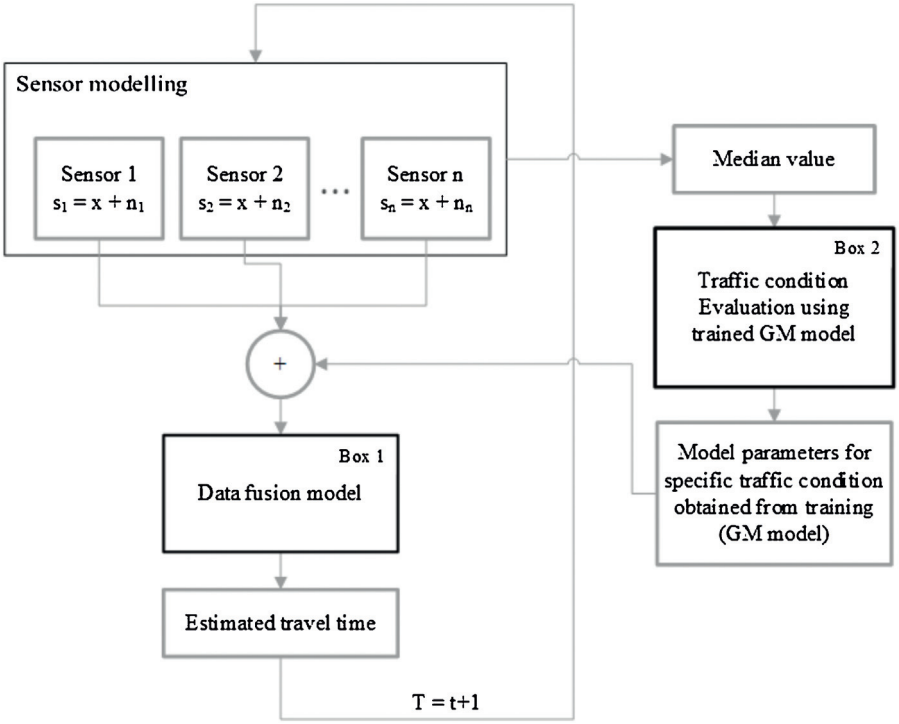


Fig. 1. Online operation scheme of the proposed model.

Where: $P(S = s|X = x)$ = likelihood of sensor (s) given the true state (x).

$P(X = x)$ = prior probability of the true state (x).

$P(S = s)$ = marginal probability of the sensor (s) under all possibilities.

Important ingredients for the Bayesian approach are the likelihood functions $P(S = s|X = x)$ and the prior distribution $P(X = x)$. Likelihood function presents the information about the parameters (x), obtained from the observed data (s), while the prior distribution presents the knowledge we have about the parameters before any measurements are made. Since the denominator does not depend on the state, the true state is estimated by maximizing the posterior distribution probability, called *maximum a posteriori* (MAP) given by:

$$x_{\mathrm{MAP}} = \arg\max p(X = x|\, S = s) \propto \arg\max\, p(S = s|\, X = x)\, P(X = x) \tag{2}$$

The non-informative or "flat" a priori probability is chosen in this study. It is assumed to have a value equal to one ($P(X = x) = 1$). This assumption is appropriate to the situation when we have no knowledge of the true distribution of the state we are estimating [38]. However, this simple assumption can be improved when we have enough data to estimate the distribution parameters of the state.

4.1. Fusion model 1: basic Bayesian data fusion approach without consideration of outliers (spurious data) – NSP model

The likelihood function or sensor model refers to the modelling of the value from a sensor given the true state. In general, it is assumed that the value observed from sensor (s) includes the true state (x) and the Gaussian noise (n):

$$s = x + n \tag{3}$$

The Gaussian noise refers to the accuracy (mean of error) and precision of estimation (variance). It is important to note that most studies found in the literature assume a zero-mean Gaussian distribution and also some studies have not clearly stated its properties [13,29]. However, a zero-mean Gaussian distribution assumption is not valid during congested period of traffic where some estimations may keep underestimating or overestimating the travel time because of assumptions in the calculation process and the nature of data. In this study, we propose a non-zero mean Gaussian distribution which can be learned from the training dataset.

The mean of errors which denotes the accuracy of estimation of a specific sensor is also called a bias of estimation. Variance or variability of estimation is considered as the precision of estimation. For example, a small positive value of the mean of errors and a small value of the variance of errors may show that the estimator is a little biased toward overestimation of the travel time, with high precision. At the training stage, these model parameters (μ_n, σ_n) can be learnt by fitting the distribution of errors to the training data set. The above assumption leads to defining the probability distribution of errors as:

$$n = s - x \quad , \text{where } n \sim N(\mu_n, \sigma_n) \tag{4}$$

$$\begin{aligned} P(S = s|X = x) &= P(S = x + n|X = x) \\ &= \frac{1}{\sqrt{2\pi}\sigma_n} e^{\frac{-((s-\mu_n)+x)^2}{2\sigma_n^2}} \end{aligned} \tag{5}$$

The above error probability distributions of individual types of sensors can be determined by a training data set. A distribution with large values of mean and variance shows less reliability in estimation and vice versa. Finally, an estimate of the travel time derived

from K types of sensors using the Bayesian data fusion approach can be written as:

$$\begin{aligned}\hat{x}_{MAP} &= argmax \prod_{k=1}^{K} P(S=s_k|X=x) \\ &= argmax \left\{ \left(\prod_{k=1}^{K} \frac{1}{\sqrt{2\pi}\sigma_k} \right) e^{\sum_{k=1}^{K} \left\{ \frac{-(s-(x+\mu_k))^2}{2\sigma_k^2} \right\}} \right\} \end{aligned} \quad (6)$$

For example, if three types of sensors are used, then the standard deviation (σ') and the mean value of the estimated travel time ($\hat{x}_{MAP}$) are obtained by setting the first partial derivative of the likelihood function, with respect to each parameter, to zero.

$$\left(\sigma'\right)^2 = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \frac{1}{\sigma_3^2} \right)^{-1} \quad (7)$$

$$\hat{x}_{MAP} = \left(\sigma'\right)^2 \left[\frac{s_1-\mu_1}{(\sigma_1)^2} + \frac{s_2-\mu_2}{(\sigma_2)^2} + \frac{s_3-\mu_3}{(\sigma_3)^2} \right] \quad (8)$$

4.2. Fusion model 2: Bayesian data fusion approach with consideration of outliers (spurious data) – SP model

Good fusion model should be able to handle properly the unconformity of data which is the common problem in data fusion. In this study, spurious model is designed specifically for this purpose. The approach has a mechanism to identify the adversarial situation and then dynamic weights are assigned based on it distance to the group (Eq. 13).

In the case of travel time data, outliers (conformity problem) may occur due to some vehicles which run very fast or very slow, compared to average vehicles. Another possible cause of outliers is due to fluctuation of traffic conditions especially during congestion. Thus, outliers should be detected and properly handled to keep the estimation at an acceptable level of accuracy. The basic Bayesian data fusion approach does not take into account outliers since only the reliability of sensors in general is considered. The probability that the estimated travel time at any instance is not spurious (r = 0), given the ground truth travel time (x) and estimated travel time from sensor (s), is written as:

$$P(r=0|X=x,S=s) = \frac{P(r=0)\,P(S=s \mid X=x, r=0)}{\sum_r (P(r)\,P(S=s|X=x,r))} \quad (9)$$

By marginalizing over the distribution of the variable r in the denominator, Eq. (10) becomes:

$$P\,(S=s|X=x) = \frac{P(r=0)P(S=s|X=x,r=0)}{P(r=0|X=x,S=s)} \quad (10)$$

Kumar et al. [29] assumed the probability of the estimation being spurious is the normal distribution with zero mean. In this study, we assumed that the probability of the estimation being spurious is the normal distribution with the bias of estimation (μ).

$$P\,(r=0|X=x,S=s) = e^{-\left\{ \frac{-((s+\mu)-x)^2}{a^2} \right\}} \quad (11)$$

This probability density function is close to 1 when the estimation ($s+\mu$) is close to the ground truth (x), and decreases exponentially when the estimation ($s+\mu$) stays away from the ground truth (x). The denominator (a) is the rate of change and conditions for the variance of sensors and distances to the other measurements. In the case of multiple sensors,

$$\begin{aligned} a_k^2 &= \frac{b_k^2}{\prod_{l\neq k,k=1}^{K} (s_k - s_l)^2} \\ b_k^2 &\geq 2\sigma_k^2 \prod_{l\neq k,k=1}^{K} (s_k - s_l)^2 \end{aligned} \quad (12)$$

Finally, the overall formula for model 2 is written as:

$$\begin{aligned} P(X=x \mid S=s_1,s_2,\ldots,s_K) &= \frac{P(X=x)}{P(S=s_1,s_2,\ldots,s_K)} \\ \times \prod_{k=1}^{K} [P(r=0)]_k \frac{1}{\sigma_k\sqrt{2\pi}} e^{-(s_k-(x-\mu_k))^2 \left\{ \frac{1}{2\sigma_k^2} - \frac{1}{a_k^2} \right\}} \end{aligned} \quad (13)$$

The mean value of state (x) is obtained by maximizing the posterior probability in Eq. (14) by setting the first derivative of the equation to zero. For instance, if three sensors are fused, then the standard deviation and mean values of fusion are obtained by:

$$\left(B'\right)^2 = \left(\frac{1}{B_1^2} + \frac{1}{B_2^2} + \frac{1}{B_3^2} \right)^{-1} \quad (14)$$

$$\hat{x}_{MAP} = \left(B'\right)^2 \left[\frac{s_1-\mu_1}{B_1^2} + \frac{s_2-\mu_2}{B_2^2} + \frac{s_3-\mu_3}{B_3^2} \right] \quad (15)$$

$$, where \;\; B_k = \frac{1}{2\left(\frac{1}{2\sigma_k^2} - \frac{1}{a_k^2}\right)}$$

The particular difference in applying the Bayesian model among studies is how sensors are modelled. It is important to denote the characteristics of sensors in the estimation of output, in this case, travel time. In most cases, the estimations are time-independent if they are applied to the object detection field of research. However, this truth does not hold with travel time estimation since travel time is spatial-temporal-dependent. Hence, in model building, it is extremely important to consider different models in term of their parameters during different traffic conditions.

5. Gaussian mixture (GM) model for traffic condition classification (Box 2)

Travel time distribution is generally used to represent travel time reliability ([34,35]). Travel time reliability indices commonly depend on the percentile of travel time which is quantified from the probability distribution model, for instance, Planning Time (95th percentile travel time over free-flow travel time). Various distribution models have been used to fit travel time data, including a lognormal distribution, a Gamma distribution, a Weibull distribution, and the Burr XII distribution [34]. Yang et al. [36] presented a modified GM model for travel time estimation. Other applications of travel time distribution were travel time prediction and energy/emission estimation from traffic [37]. In a similar manner, this study uses travel time distribution to classify traffic conditions. In different traffic conditions, different sources of traffic data have different levels of accuracy. This may result in inconsistency and sometimes even contradictory fusion results if the fusion model does not take traffic conditions into account. So it is necessary to classify the traffic conditions which lead to different parameters of the fusion model (mean and variance). In classifying traffic conditions, the traditional method uses the level of service (LOS), according to the Highway Capacity Manual (HCM)

[43]. However, calibration is required when it is applied to different types of flow. Arabani et Pourzeynali [39] proposed a methodology based on fuzzy logic. Speed, maximum service volume, and the ratio of volume/capacity and density were used as inputs. A similar methodology proposed by [13] used flow, speed, and density to classify the traffic condition into free flow (LOS: A, B, C) and congested flow (LOS: F). However, this methodology requires many types of input in order to classify effectively. In this study, the Gaussian mixture (GM) model is used. The GM model is a probabilistic model that all data points are assumed to come from a mixed finite Gaussian distribution with unknown parameters. It is a widely used statistical method for clustering. On the other hand, in traffic engineering, it is commonly used to model the distribution of traffic variables, such as flow, occupancy in [40]. Even though, there have been a limited number of studies using the GM model for a traffic classifier. To apply this clustering technique, predefined numbers of components to be clustered is required, corresponding to the number of traffic conditions. Unlike LOS, the traffic conditions are usually divided into only two or at most three conditions [41]. The probability density function of the univariate Gaussian mixture model is the weighted average of the normal probability density functions of different traffic conditions ($N\left(X|\mu_j, \sigma_j\right)$) as shown in the following equation.

$$P(X) = \sum_{j=1}^{J} \pi_j N(X \mid \mu_j, \sigma_j) \tag{16}$$

The expected maximization algorithm (EM) is used in this study to fit the parameters of the model (mean (μ_j), variance (σ_j), and mixing coefficient (π_j)). Details of the EM algorithm can be found in [42].

In this study, two clusters (free flow and congested flow) and three clusters (free flow, congested flow, and transitional flow) are proposed and compared. GM model is presented in Box 2 (Fig. 1) where the traffic data input from sensor is evaluated for associated state and properties (means and variances) obtained during training stage.

6. Case studies

In the following sub-sections, three case studies are investigated. The performance measurements of the proposed models are compared among different fusion models and different traffic assumptions.

6.1. A simulated case study

This section presents the case study using simulated data. The study site is around 5 km, stretched in the northbound direction of freeway I-880, between the intersection of Stevenson Boulevard and Thornton Ave, in Alameda County, California (Fig. 2). The study corridor consists of 6 on-ramps and 2 off-ramps denoted by the crosslines, in which the exact locations are presented by the distances at the top of the figure. The simulation inputs are the real-world flow data retrieved from the website of the Caltrans Performance Measurement System (PeMS)[1]. The time dependent traffic flow data on March 10, 2014 at the on-ramps and off-ramps of the network are used to set up the simulation.

The microscopic simulation model is performed using AIMSUN 8.1. Multiple simulations (1000 simulations) are performed in order to obtain the average result during 16 h from 6:00 AM to 10:00 PM. The travel time data are derived from the simulated sensors which are placed along the road. Multiple types of simulated sensors are: loop detector, GPS, and virtual trip line (VTLs[2]). The mid-block loop detectors are placed for all links to collect speed. Similarly, the VTLs are placed to collect the speed of probe vehicles, while the 5% trajectory data is extracted from all vehicles, represented as GPS data. By referencing the ground truth travel time, the measurements of errors are compared among different proposed models. Fig. 3 shows the estimated travel time calculated from each type of sensor during the simulations. The training stage is divided into two steps: fitting the GM model using Eq. 16 (Figs. 4 and 5) and fitting estimated errors to the normal distribution function using Eq. 5 (Table 1). For this case study, 1000 times of bootstrap sampling, consisting of 50% of the total data for training and the remaining 50% for validation are used (Fig. 6).

The training results show that, for the GM model with two components, the mean travel time during free flow (the first component) is 322 s with a standard deviation of 27 s During congested flow (the second component), the mean travel time is 498 s with a standard deviation of 120 s. The components of traffic conditions consist of 60% free flow and 40% congested flow. For the GM model with three components, the mean travel time during free flow (the first component) is around 320 s with a standard deviation of 25 s. During transitional flow (the second component), the mean travel time is 437 s with a standard deviation of 64 s. During congested flow (the third component), the mean travel time is 678 s with standard deviation of 44 s. The components of traffic conditions consist of 60% free flow, 30% transitional flow, and 10% congested flow condition accordingly. It is essential to note that, the training stage in this study is performed offline, however, it is also possible to perform it in real time.

Table 1 presents the results of the characteristics of errors of the estimation for each type of sensor. It is observed that GPS sensors overestimate the travel time since the mean of errors is positive. Similarly, VTLs and loop detectors underestimate the travel time as the means of errors are negative. Furthermore, the standard deviation of the travel time estimated by GPS sensors is the largest, followed by those estimated from loop detector sensors and VTL sensors. In other words, GPS sensors provide the least precise estimation, followed by loop detector sensors and VTL sensors. In general, the quality of travel time estimation from different types of sensors give different levels of errors due to different factors.

It is of interest to compare the errors between the proposed model and baseline models. In this study the simple fusion models (mean and median) are selected as the baseline models. There is a number of comparison indices used in the literature, however, the common statistical measurement, the mean absolute percentage error (MAPE)[3], is use as the common measurement to compare among different case studies. Even though other measurements of accuracy are also given for references including Mean Absolute Error (MAE), Mean Absolute Scale Error (MASE) and Mean Signed Difference (MSD).

In order to assess the performance of the proposed fusion models, 1000 times of bootstrap sampling of the training and validating data (50% each) are used. The average result (Table 2) shows the improvement of the proposed models (NSP and SP models) compared to the baseline fusion models. The MAPE decreases from 4.30% to 3.44% when compared with the best result of the base-

[1] http://pems.dot.ca.gov/

[2] Virtual Trip Lines are geographic markers where updated locations of probe vehicles are provided to avoid privacy sensitive locations [43].

[3] The mean absolute percentage is calculated by $MAPE = \left(\sum_{i=1}^{n} \left|e_i/T_i\right| \times 100\right)/n$, where e is a column vector of estimation errors, and T is a column vector of n observed values.

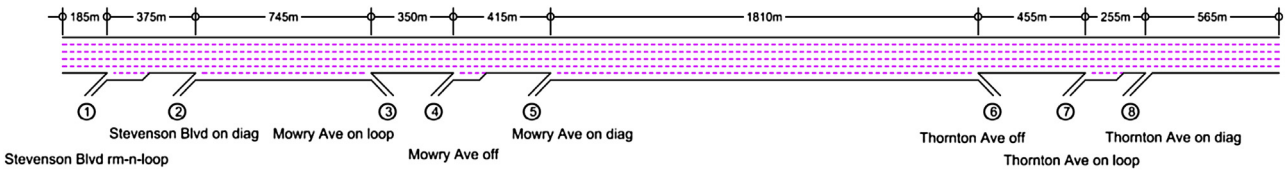


Fig. 2. Schematic figure of simulated case study site.

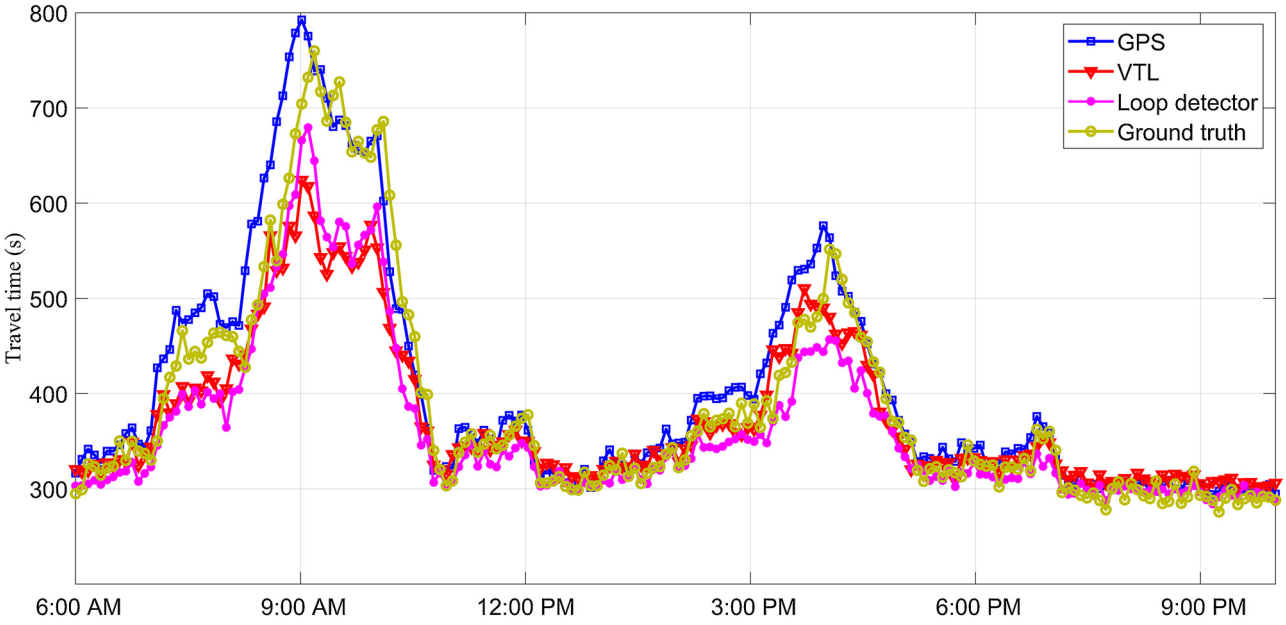


Fig. 3. Estimated travel time from different sensors.

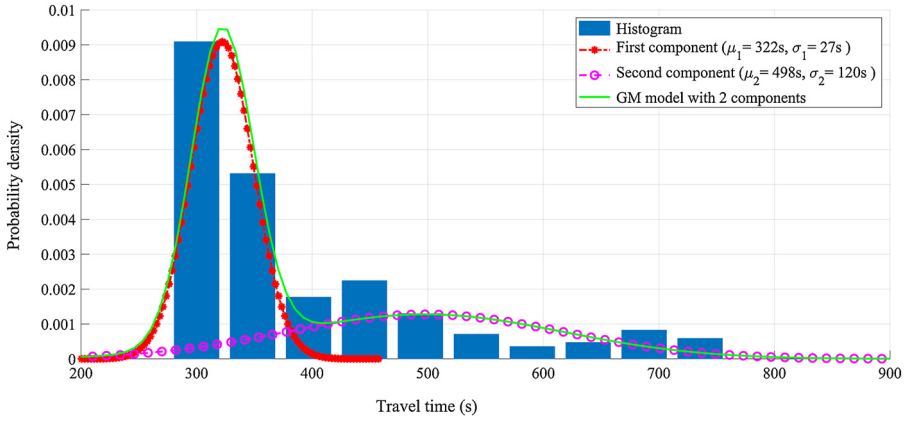


Fig. 4. GM model with two components of traffic.

Table 1
Error of estimation by each type of sensor: simulated case study.

Travel time estimation based on	MAPE (%)									
	Two components				Three components					
	Free flow		Congested flow		Free flow		Transitional flow		Congested flow	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
GPS	12.5*	10.2	20.8*	47.1	11.5*	10.4	30.3*	41.9	22.9*	47.8
Loop detector	−10.5*	11.9	−62.4*	25.4	6.1*	11.1	−19.9*	28.9	−132*	33.4
VTLs	7.8*	11.6	−35.5*	37.8	−5.8*	11.2	−40.5*	22	−104*	32.5

Note: * indicates that mean errors are statically different from zero, at the 95% confidence level.

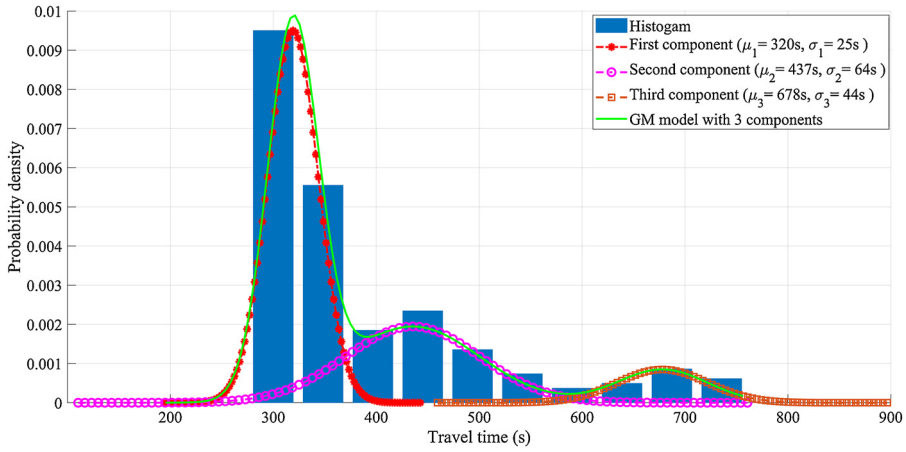
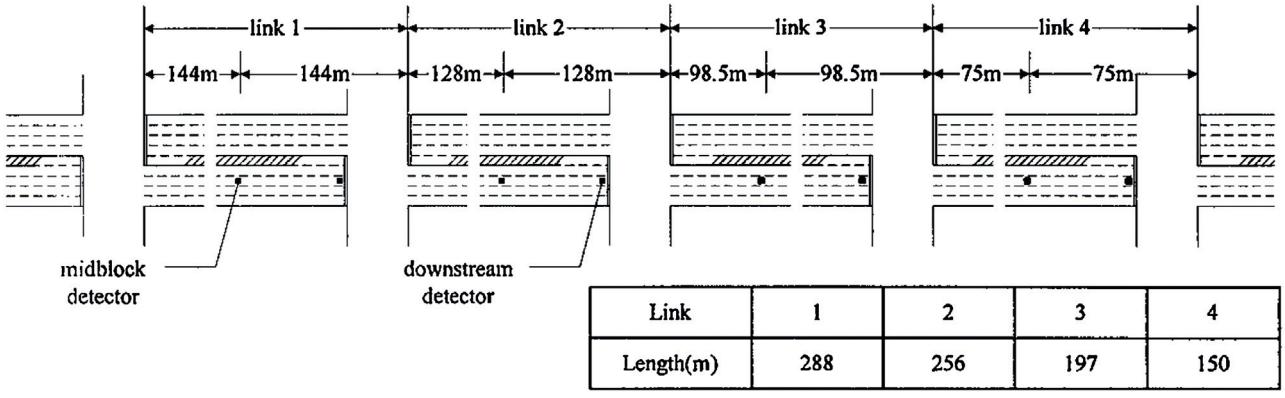


Fig. 5. GM model with three components of traffic.



Link	1	2	3	4
Length(m)	288	256	197	150

Fig. 6. Schematic figure of case study.

(Source: [31])

line model (Arithmetic mean) and NSP model with three traffic conditions. The improvement is approximately 20% of the baseline model's MAPE. However, no significant improvement is found between the NSP model and SP model. This is because the data from simulation is not so spurious (containing many outliers), compared with real-world conditions. A relative improvement of 10% is also found when the model with three traffic conditions is used instead of two traffic conditions due to better classification of the state of traffic. Motivated by this interesting result, more applications of the proposed model on real-world data are presented in the next section.

6.2. Real-world case studies

To investigate the applicability of the model in real-world applications, two real-world case studies are conducted. The first case study utilizes the data provided by [31], while the second case study employs the experimental data from the Mobile Century Project [30].

6.2.1. First case study: Suwon arterial road

The dataset in this case study was collected from the Suwon arterial road in South Korea for a periods of 2 h (11:00 to 13:00) on November 21 and 28, 1998. The selected road section consists of four links with an average length around 200 m. Travel time data was aggregated into 5 min intervals. The calculated travel time derived from an individual type of sensor was presented in Choi and Chung (2002). The data consist of loop detector data, GPS data, historical data, and ground truth data. The tabulated travel time data derived from the individual type of sensor used in this case study can be found in [32].

As the available data are limited, in order to get an unbiased estimation of parameters, 1000 times of bootstrap samplings are carried out. For each sampling, 80% of the data are randomly selected and used for the training while the rest (20% of the data) is used for validation. Since only two hours of data was available during non-peak hours, only one traffic condition is considered in the modelling. In this case, fitting of the probability density function of the error (Eq. 5) can be performed immediately. It can be seen from the result (Table 3) that the estimation using loop detector data is the most precise estimation due to its small variance (link 1, link 2, and link 3). On the other hand, travel time estimated from GPS data shows the lowest precision while using historical travel time data provides moderately precise estimation. For the bias of the estimation, using only historical travel time data seems to overestimate the travel time while using only loop detector or GPS data seems to underestimate the travel time. Overall, the results show that the quality of estimation from each sensor may give different errors on different links due to different factors, such as network characteristics and traffic conditions.

Table 2
Performance of the fusion models: simulated case study.

(a) Models with two traffic conditions

Estimation models		Free Flow		Congested Flow		Overall
	Mean	Std.	Mean	Std.	Mean	
		MAPE (%)				
Baseline models	**Arithmetic mean**	2.95	2.34	7.01	4.92	4.3
	Median	3.15	2.66	8.99	5.83	5.1
Bayesian approach	**NSP Model**	2.71*	0.39	5.99*	0.99	3.81
	SP Model	2.80*	0.44	6.21*	0.99	3.94
		MAE(s)				
Baseline models	**Arithmetic mean**	9.46	7.71	39.31	33.71	19.41
	Median	10.12	8.94	50.7	41	23.65
Bayesian approach	**NSP Model**	8.88*	1.31	32.4*	5.97	16.72
	SP Model	9.2*	1.49	33.2*	6.18	17.2
		MASE				
Baseline models	**Arithmetic mean**	0.82	0.67	1.69	1.43	1.11
	Median	0.87	0.78	2.18	1.74	1.3
Bayesian approach	**NSP Model**	0.5*	0.08	0.5*	0.11	0.5
	SP Model	0.52*	0.09	0.5*	0.11	0.51
		MSD(s)				
Baseline models	**Arithmetic mean**	−3.9	N/A	30.77	N/A	7.65
	Median	−3.65	N/A	45.75	N/A	12.81
Bayesian approach	**NSP Model**	−0.28	N/A	1.5	N/A	0.31
	SP Model	0.7	N/A	3.2	N/A	1.53

(b) Models with three traffic conditions

Estimation models		Free Flow		Transitional Flow		Congested Flow		Overall
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	
			MAPE (%)					
Baseline models	**Arithmetic mean**	2.89	2.13	5.41	3.87	10.41	5.95	4.3
	Median	2.99	2.24	6.69	4.66	14.32	5.6	5.1
Bayesian approach	**NSP Model**	2.54*	0.38	4.65*	0.93	5.69*	1.69	3.44
	SP Model	2.59*	0.38	4.87*	1.03	5.39*	1.52	3.51
			MAE(s)					
Baseline models	**Arithmetic mean**	9.12	6.62	24.38	3.87	71.34	41.1	19.25
	Median	9.43	6.94	29.8	23.06	97.66	39.75	23.43
Bayesian approach	**NSP Model**	8.14*	1.23	21.39*	4.7	37.9*	10.75	14.65
	SP Model	8.28*	1.27	22.3*	4.98	36.7*	9.65	14.88
			MASE					
Baseline models	**Arithmetic mean**	0.78	0.57	1.35	1.04	2.6	1.42	1.11
	Median	0.81	0.6	1.65	1.23	3.46	1.37	1.29
Bayesian approach	**NSP Model**	0.48	0.08	0.47	0.13	0.95	0.48	470
	SP Model	0.49	0.09	0.49	0.15	0.93	0.48	467
			MSD(s)					
Baseline models	**Arithmetic mean**	−4.41	N/A	13.19	N/A	69.56	N/A	7.47
	Median	−4.28	N/A	21.56	N/A	97.66	N/A	12.54
Bayesian approach	**NSP Model**	0.14	N/A	0.18	N/A	−0.6	N/A	0.08
	SP Model	1.12	N/A	−0.4	N/A	−0.2	N/A	0.56

Note: * the estimated error is statistically different from the baseline model (median) at the 95% confidence level.

Table 3
Error of estimation by each type of sensor: Suwon arterial road.

Travel time estimation based on	Link 1		Link 2		Link 3		Link 4	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Historical travel time	7.82*	15.1	3.48*	5.38	7.74*	10.0	4.15*	4.54
Loop detector data	−16.3*	11.40	0.35*	1.32	−2.95*	4.06	−10.10*	12.23
GPS data	3.73	35.3	−4.34*	4.09	−27.35*	35	−39.05*	55.74

Note: * mean value is statistically different from zero at the 95% confidence level.

The proposed models outperform the baseline models and the method proposed by [31]. The MAPE reduces from 20.92% to 17.5% for the SP model, which corresponds to approximately 16.3% improvement, compared to the model proposed by Choi and Chung (2002). In this case study, the NSP model and SP model provided similar accuracy (Table 4) because the traffic condition during the study period was relatively uncongested with small number of outliers (non-spurious data).

6.2.2. Second case study: I800 N corridor

Open source data provided under the Mobile Millennium Project are used for this case study. The data consist of loop detector data, VTLs data, GPS data, and ground truth data. They were collected on February 08, 2008 from 10:00 until 18:00 along the I880 N corridor in California. The 16 km long section between Decoto Road to Winton Avenue is selected for this case study. As the data is provided in raw format, data pre-processing is necessary in order to obtain travel time estimated from each type of sensor (Fig. 7). The

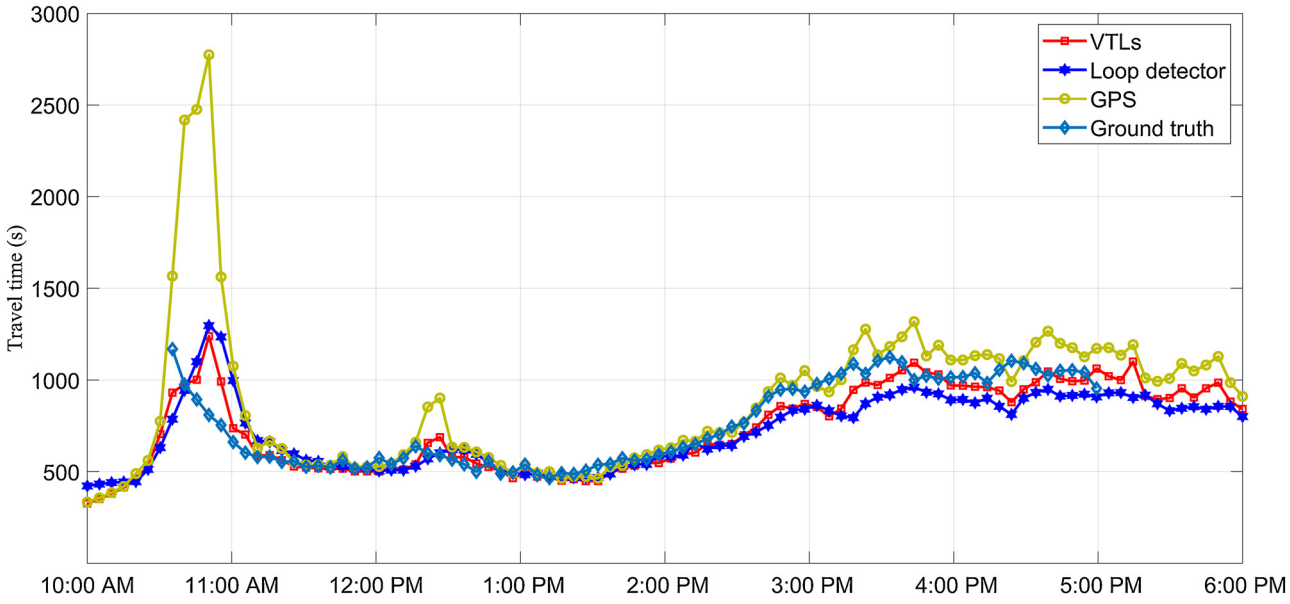


Fig. 7. Estimated travel time from different sensors.

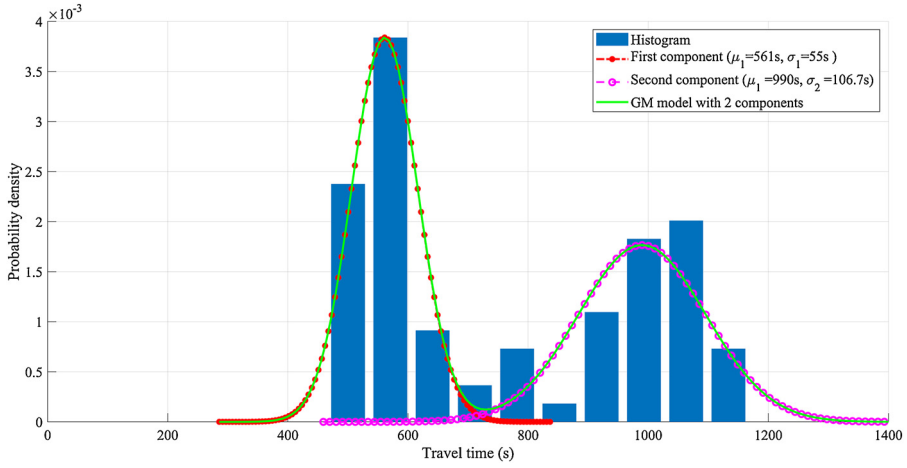


Fig. 8. GM model with two components of traffic.

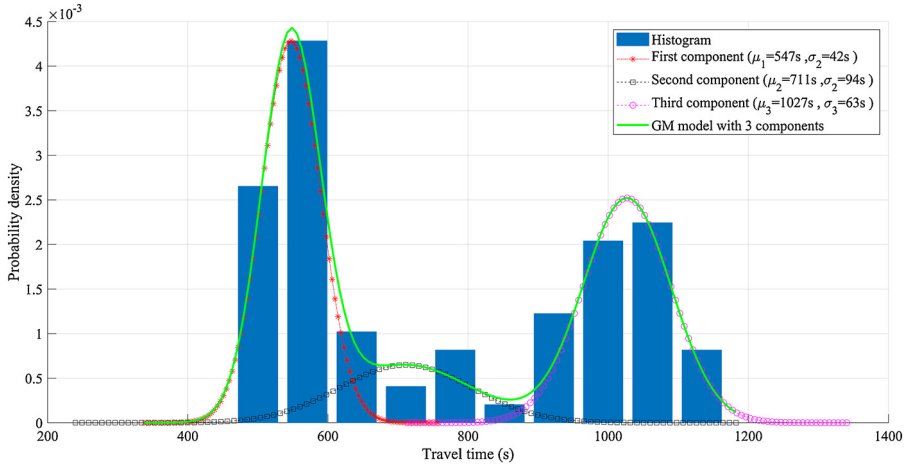


Fig. 9. GM model with three components of traffic.

Table 4
Performance of the fusion models: Suwon arterial road.

Estimation models		Link 1		Link 2		Link 3		Link 4		Overall
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	
					MAPE (%)					
Baseline models	Arithmetic mean	37	32.51	17	11.35	40	38.33	26	22.78	30
	Median	30	22.72	15	11.95	24	20.85	15	10.39	21
	Choi and Chung (2002)	26.51	N/A	16.48	N/A	15.16	N/A	25.55	N/A	20.92
Bayesian approach	NSP model	21.1*	8.8	18.3*	5.9	12.9*	4.6*	18.3*	6.8*	17.7
	SP model	21.8*	8.6	17.8*	5.4	13.3*	4.7*	17.1*	5.0*	17.5
					MAE(s)					
Baseline models	**Arithmetic mean**	15.52	12	5.02	5.51	12.84	14.71	8.76	8.69	10.53
	Median	13.73	12.76	4.57	5.6	7.95	9.09	4.47	3.9	7.68
	Choi and Chung (2002)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Bayesian approach	**NSP model**	9.15*	3.92	5.68*	2.78	4*	2.13	5*	2.42	5.95
	SP model	10.62*	4.91	5.03*	2.34	3.85*	1.63	4.8*	1.92	6.075
					MASE					
Baseline models	**Arithmetic mean**	0.93	0.71	0.67	0.73	1.93	2.15	0.87	0.84	1.1
	Median	0.82	0.76	0.61	0.74	1.18	1.33	0.44	0.37	0.76
	Choi and Chung (2002)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Bayesian approach	**NSP model**	0.54*	0.28	0.72*	0.37	0.62*	0.3	0.5*	0.3	0.59
	SP model	0.61*	0.34	0.65	0.36	0.63*	0.37	0.48*	0.2	0.59
					MSD(s)					
Baseline models	**Arithmetic mean**	−0.31	N/A	1.55	N/A	−10.5	N/A	−2.9	N/A	−3.04
	Median	5.56	N/A	2.4	N/A	−5.12	N/A	2.83	N/A	1.41
	Choi and Chung (2002)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Bayesian approach	**NSP model**	0.6	N/A	0.62	N/A	0.15	N/A	0.31	N/A	0.42
	SP model	2.36	N/A	1.51	N/A	0.3	N/A	1.2	N/A	1.34

Note: * estimated error is statistically different from the baseline model (median) at the 95% confidence level.

Table 5
Error of estimation for each type of sensor: I800 N corridor.

Travel time estimation based on	MAPE (%)									
	Two components				Three components					
	Free flow		Congested flow		Free flow		Transitional flow		Congested flow	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
GPS	26*	27.4	136*	345	14.9*	35	271.5*	380	109.6*	117
Loop detector	9.77	45.8	−82*	201	13.2*	58	−68.6*	5.49	−142*	44.6
VTLs	−14*	38.4	−48*	109	−20.6*	33.1	−61.7*	28.5	−46.2*	69

Note: * mean value is statistically different from zero at the 95% confidence level.

details of methods and procedures of calculation can be found in [32]. In this case study, 80% of the data are randomly selected for GM model training. Figs. 8 and 9 show the results of the classification models according to two and three components of traffic. Furthermore, Table 5 presents the results of the characteristics of errors of the estimation by each type of sensor. It is clearly seen that different types of sensors give different levels of errors due to different factors such as traffic condition and types of sensors. In this case study, during free flow, loop detectors give the most accurate travel time estimation. On the other hand, during congested flow, the most accurate travel time estimation can be obtained from VTLs.

Interestingly, in this case study, it is found that the SP model performs better than the NSP model (Table 6) in most cases. In overall traffic conditions, the value of the MAPE reduces moderately from 8.35% to 7.95% (for the two component GM model) and 9.37% to 8.50% (for the three component GM model) when SP models are used instead of NSP models. In particular, significant improvement of the SP model over the NSP model are found during congested flow. Overall, the NSP model performs slightly better in free flow, while the SP model performs better in transitional and congested flow. Thus, it is recommended by the authors to apply the NSP-SP model which combines the NSP and SP models. In the NSP-SP model, the NSP model is applied in free flow condition and the SP model is applied in transitional and congested flow. The average improvement of model performance of the proposed NSP-SP model over the benchmark model (median) is approximately 17%. In this case study, no improvement is found between the assumption of two conditions of traffic (7.67%) and three conditions of traffic (8.46%). This finding contradicts the previous conclusion made in the first case study. To the authors' knowledge, this effect may come from the amount of data which is not statistically sufficient to get an unbiased estimator of model parameters for the three components of traffic assumption, especially during transitional flow.

7. Conclusions

This study proposes a framework for travel time estimation by combining the modified Bayesian data fusion and Gaussian mixture model for different traffic conditions. Three case studies are presented. Data utilized include, but are not limited to, three types of sensors. Significant improvement, at the 95% level of confidence for an accuracy of at least 16.3% over the baseline model's MAPE value, is found. This improvement also proved by others measurements of accuracy (MAE, MASE, MSD) given for references. Besides the improvement of the accuracy, the model also provided higher reliability of estimation due to its lower variances. In every case study, it was found that variances of travel time estimation are much reduced. Especially, in the second real-world case study, the standard deviation of the MAPE was greatly reduced, from 6.07% to 2.53%. Higher reliability of the model is preferable in real-world applications because it gives more reliable estimates to users.

Table 6
Performance of the fusion models: I800 N corridor.

(a) Two components of traffic

Estimation models		MAPE (%)				
		Free flow		Congested flow		Overall
		Mean	Std.	Mean	Std.	
Baseline models	**Arithmetic mean**	7.27	8.14	14.35	24.25	10.59
	Median	8.17	9	12.25	13.48	10.08
Bayesian approach	**NSP model**	7.15*	2.18	10.35*	5.69	8.35
	SP model	7.60*	2.74	8.54*	5.2	7.95
	NSP-SP model	7.15*	2.18	8.54*	5.2	7.67
		MAE(s)				
Baseline models	**Arithmetic mean**	42.1	51.19	130	200	75.06
	Median	47.65	57.51	115.31	108.22	73.02
Bayesian approach	**NSP model**	41.33*	14.14	97.27*	48.37	62.31
	SP model	44.77	18.15	90.05*	41.6	61.75
	NSP-SP model	41.33*	14.14	90.05*	41.6	59.60
		MASE				
Baseline models	**Arithmetic mean**	1.54	1.87	3.05	4.61	2.10
	Median	1.74	2.1	2.57	2.49	2.05
Bayesian approach	**NSP model**	0.92*	0.37	1.28*	0.62	1.06
	SP model	0.98*	0.41	1.15*	0.57	1.04
	NSP-SP model	0.92*	0.37	1.15*	0.57	1.00
		MSD(s)				
Baseline models	**Arithmetic mean**	−11.02	N/A	−32.25	N/A	−18.98
	Median	−8.02	N/A	35.25	N/A	8.21
Bayesian approach	**NSP model**	−6	N/A	−23.27	N/A	−12.48
	SP model	−7.9	N/A	−20.9	N/A	−12.78
	NSP-SP model	−6	N/A	−20.9	N/A	−11.59

(b) Three components of traffic

Estimation models		MAPE(%)						
		Free flow		Transitional flow		Congested flow		Overall
		Mean	Std.	Mean	Std.	Mean	Std.	
Baseline models	Arithmetic mean	6.62	6.31	27	38.54	9.89	14.33	10.59
	Median	7.1	6.04	23.56	24.44	9.18	6.07	10.08
Bayesian approach	NSP model	6.83*	2.07	26.99*	23.55	8.20*	4.27	9.37
	SP model	6.91*	1.92	22.86*	13.18	6.70*	2.53	8.5
	NSP-SP model	6.83*	1.92	22.86*	13.18	6.70*	2.53	8.46
		MAE(s)						
Baseline models	**Arithmetic mean**	36.92	36.82	205.31	307.81	98.38	132.28	74.3
	Median	39.68	34.87	174.74	186.08	94.5	63.51	71.45
Bayesian approach	**NSP model**	37.28*	11.44	197.55*	174.03	81.06*	39.1	68.11
	SP model	38.03	11.14	169.47	105.08	67.83*	23.95	61.34
	NSP-SP model	37.28*	11.44	169.47	105.08	67.83*	23.95	60.91
		MASE						
Baseline models	**Arithmetic mean**	1.45	1.45	3	7.61	2.43	3.24	1.92
	Median	1.56	1.37	3.46	4.6	2.2	1.55	1.96
Bayesian approach	**NSP model**	1*	0.46	4.17*	5.58	1.46*	0.85	1.48
	SP model	1.03*	0.47	3.86*	4.82	1.17*	0.49	1.37
	NSP-SP model	1*	0.46	3.86*	4.82	1.17*	0.49	1.35
		MSD(s)						
Baseline models	**Arithmetic mean**	−8.3	N/A	−143.44	N/A	3.89	N/A	−18.44
	Median	−4.83	6.04	−86.93	N/A	63.86	N/A	8.8
Bayesian approach	**NSP model**	−5.2	N/A	−91	N/A	−19.5	N/A	−10.22
	SP model	−4.3	N/A	−62.89	N/A	−6.98	N/A	−11.26
	NSP-SP model	−5.2	N/A	−62.89	N/A	−6.98	N/A	−11.78

Note: * estimated errors is statistically different from the baseline model (median) at the 95% confidence level.

It is important to note that the historical value of travel time plays an important role in improving the accuracy of the travel time estimation. However, despite a small amount of data available in the case studies, the performance of the proposed model is highly satisfactory. Even though a general framework of offline learning is proposed in this study, it is also possible to apply online learning for model parameters which may lead to better estimation for sudden changes in traffic conditions such as the case of traffic accidents.

This study provides two significant contributions to travel time estimation research based on the Bayesian data fusion approach. Two additional features were put into existing models including the difference of traffic conditions classified by the Gaussian mixture (GM) model (Section 5) and the bias estimation from individual sensor by introducing a non-zero mean Gaussian distribution which parameters are learnt from the training dataset (Section 4.1). The different conditions of traffic information were taken into account by the Gaussian mixture model so that the proposed model would be supplied with appropriate parameters at specific traffic condition. The non-zero mean Gaussian distribution helps to reduce the biases of sensors that usually occur during transitional and congested flow. From case studies presented, it was found that spurious data do exist frequently. Understanding this phenomena,

better estimation was achieved in this study by combining our non-spurious and spurious models.

It is noted that the case studies provided in this study rely on the link and corridor (instantaneous sum of all link travel time) based travel time estimation. Using other estimation techniques, such as the time slide model [44], dynamic time slice model [45], linear model [46], and trajectory reconstruction model [47], may further improve our estimation. In addition, our model does not consider delay time, which may occur at an intersection, mostly appearing in urban areas. The travel time delay at any considered node can be learned and applied in the model such as found in [48]. These issues are recommended for further research. Finally, the proposed model gives noticeable high accuracy of estimation with low computational effort, and could be applied with existing traffic measurement technologies available. Moreover, we believe the proposed model could be adopted in any areas of sensor data fusion where errors adjustment based on the conditional of fusion environment, may involve i.e. obstacle detection for autonomous vehicle, multi-target tracking from different sensor such as camera, radar, and lidar.

Acknowledgement

This research is financially supported by Sirindhorn International Institute of Technology (SIIT), Thammasat University. The authors would like to thank the six anonymous reviewers and the editor for their valuable comments and suggestions that greatly contributed to improving the final version of the paper.

References

[1] X.J. Ban, R. Herring, J.D. Margulici, A.M. Bayen, Optimal sensor placement for freeway travel time estimation, in: Transportation and Traffic Theory 2009: Golden Jubilee, Springer, 2009, pp. 697–721.
[2] J. Palen, The need for surveillance in intelligent transportation systems, Part Two Intellimotion 6 (2) (1997).
[3] D. Ettema, H. Timmermans, Costs of travel time uncertainty and benefits of travel time information: conceptual model and numerical examples, Transp. Res. Part C Emerg. Technol. 14 (2006) 335–350, http://dx.doi.org/10.1016/j.trc.2006.09.001.
[4] L. Chu, S. Oh, W. Recker, Adaptive Kalman Filter Based Freeway Travel Time Estimation, 2005.
[5] J.-S. Oh, R. Jayakrishnan, W. Recker, Section Travel Time Estimation from Point Detection Data, Center for Traffic Simulation Studies, 2002.
[6] F. Soriguera, D. Rosas, F. Robusté, Travel time measurement in closed toll highways, Transp. Res. Part B Methodol. 44 (2010) 1242–1267, http://dx.doi.org/10.1016/j.trb.2010.02.010.
[7] P. Chaudhuri, P.T. Martin, A.Z. Stevanovic, C. Zhu, The effects of detector spacing on travel time prediction on freeways, World Acad. Sci. Eng. Technol. 66 (2010) 1–10.
[8] J. Brügmann, Modelling and Implementation of a Microscopic Traffic Simulation System, Logos Verlag Berlin GmbH, 2015.
[9] H.B. Celikoglu, Flow-based freeway travel-time estimation: a comparative evaluation within dynamic path loading, Ieee Trans. Intell. Transp. Syst. 14 (2013) 772–781, http://dx.doi.org/10.1109/TITS.2012.2234455.
[10] U. Mori, A. Mendiburu, M. Álvarez, J.A. Lozano, A review of travel time estimation and forecasting for advanced traveller information systems, Transp. A Transp. Sci. 11 (2015) 119–157, http://dx.doi.org/10.1080/23249935.2014.932469.
[11] B. Hellinga, L. Fu, Assessing expected accuracy of probe vehicle travel time reports, J. Transp. Eng. 125 (1999) 524–530.
[12] K.K. Sanwal, J. Walrand, Vehicles as Probes, California Partners for Advanced Transit and Highways (PATH), 1995.
[13] F. Soriguera Martí, D. Abeijón Monjas, L. Thorson Bofarull, F. Robusté Antón, Highway Travel Time Data Fusion, Transportation Research Board, 2009.
[14] A. Tarko, N. Rouphail, Travel time data fusion in advance, Pacific Rim TransTech Conference (1993: Seattle, Wash.). Proceedings Pacific Rim TransTech Conference 1 (1993).
[15] N.-E. El Faouzi, Data-driven Aggregative Schemes for Multisource Estimation Fusion: a Road Travel Time Application, 2004, pp. 351–359, http://dx.doi.org/10.1117/12.541336.
[16] C. Bachmann, B. Abdulhai, M.J. Roorda, B. Moshiri, A comparative assessment of multi-sensor data fusion techniques for freeway traffic speed estimation using microsimulation modeling, Transp. Res. Part C Emerg. Technol. 26 (2013) 33–48, http://dx.doi.org/10.1016/j.trc.2012.07.003.
[17] C. Nanthawichit, T. Nakatsuji, H. Suzuki, Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway, Transp. Res. Rec.: J. Transp. Res. Board 1855 (2003) 49–59, http://dx.doi.org/10.3141/1855-06.
[18] T. Tettamanti, M.T. Horváth, I. Varga, Road Traffic Measurement and Related Data Fusion Methodology for Traffic Estimation, Transport and Telecommunication, 15, 2014 (Accessed 4 May 2018) https://trid.trb.org/view/1422277.
[19] A. Nantes, D. Ngoduy, A. Bhaskar, M. Miska, E. Chung, Real-time traffic state estimation in urban corridors from heterogeneous data, Transp. Res. Part C Emerg. Technol. 66 (2016) 99–118, http://dx.doi.org/10.1016/j.trc.2015.07.005.
[20] Y.-J. Byon, A. Shalaby, B. Abdulhai, S. El-Tantawy, Traffic data fusion using SCAAT kalman filters, in: Transportation Research Board 89th Annual Meeting, 2010.
[21] P. Nelson, P. Palacharla, A neural network model for data fusion in advance, Pacific Rim TransTech Conference (1993: Seattle, Wash.). Proceedings Pacific Rim TransTech Conference 1 (1993).
[22] J.N. Ivan, V. Sethi, Data fusion of fixed detector and probe vehicle data for incident detection, Computer-Aided Civil and Infrastructure Engineering. 13 (1998) 329–337.
[23] X. Du, M. El-Khamy, J. Lee, L. Davis, Fused DNN: a deep neural network fusion approach to fast and robust pedestrian detection, 2017 IEEE Winter Conference on Applications of Computer Vision (WACV) (2017) 953–961, http://dx.doi.org/10.1109/WACV.2017.111.
[24] Y. Liu, X. Chen, H. Peng, Z. Wang, Multi-focus image fusion with a deep convolutional neural network, Inf. Fusion 36 (2017) 191–207, http://dx.doi.org/10.1016/j.inffus.2016.12.001.
[25] J.V. Tu, Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes, J. Clin. Epidemiol. 49 (1996) 1225–1231, http://dx.doi.org/10.1016/S0895-4356(96)00002-9.
[26] F. Soriguera, F. Robusté, Highway travel time accurate measurement and short-term prediction using multiple data sources, Transportmetrica 7 (2011) 85–109, http://dx.doi.org/10.1080/18128600903244651.
[27] E. Faouzi, N.-C. Simon, Travel Time Estimation by Evidential Data Fusion, Recherche – Transports – Securite (French), 2000.
[28] Q.-J. Kong, Z. Li, Y. Chen, Y. Liu, An approach to urban traffic state estimation by fusing multisource information, IEEE trans. Intell. Transp. Syst. 10 (2009) 499–511, http://dx.doi.org/10.1109/TITS.2009.2026308.
[29] M. Kumar, D.P. Garg, R.A. Zachery, A method for judicious fusion of inconsistent multiple sensor data, IEEE Sens. J. 7 (2007) 723–733, http://dx.doi.org/10.1109/JSEN.2007.894905.
[30] J.C. Herrera, D.B. Work, R. Herring, X.Jeff Ban, Q. Jacobson, A.M. Bayen, Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment, Transp. Res. Part C Emerg. Technol. 18 (2010) 568–583, http://dx.doi.org/10.1016/j.trc.2009.10.006.
[31] K. Choi, Y. Chung, A data fusion algorithm for estimating link travel time, J. Intell. Transp. Syst. Technol. Plan. Oper. 7 (2002) 235–260, http://dx.doi.org/10.1080/714040818.
[32] S. Mil, M. Piantanakulchai, Travel time estimation based on fused traffic state data: case studies in US and South Korea, J. Eastern Asia Soc. Transp. Stud. 11 (2015) 1868–1884, http://dx.doi.org/10.11175/easts.11.1868.
[33] Multi-Sensor Data Fusion: An Introduction, Softcover, Springer, 2010.
[34] Y. Guessous, M. Aron, N. Bhouri, S. Cohen, Estimating travel time distribution under different traffic conditions, Transp. Res. Procedia 3 (2014) 339–348, http://dx.doi.org/10.1016/j.trpro.2014.10.014.
[35] C.R. Sekhar, Y. Asakura, Modelling travel time distribution under various uncertainties on Hanshin expressway of Japan, Eur. Transp. Res. Rev. 6 (2014) 85–92, http://dx.doi.org/10.1007/s12544-013-0111-3.
[36] Q. Yang, G. Wu, K. Boriboonsomsin, M. Barth, Arterial Roadway Travel Time Distribution Estimation and Vehicle Movement Classification Using a Modified Gaussian Mixture Model, IEEE, 2013, pp. 681–685.
[37] Q. Yang, K. Boriboonsomsin, M. Barth, Arterial roadway energy/emissions estimation using modal-based trajectory reconstruction, 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC) (2011) 809–814, http://dx.doi.org/10.1109/ITSC.2011.6083069.
[38] Default Values for Highway Capacity and Level of Service Analyses, Transportation Research Board, Washington, D.C, 2008.
[39] M. Arabani, S. Pourzeynali, Fuzzy Logic Methodology to Evaluate the Service Level of Freeways Basic Segments*, n.d.
[40] X. Liu, L. Pan, X. Sun, Real-time traffic Status classification based on gaussian mixture model, 2016 IEEE First International Conference on Data Science in Cyberspace (DSC) (2016) 573–578, http://dx.doi.org/10.1109/DSC.2016.39.
[41] B.S. Kerner, Criticism of generally accepted fundamentals and methodologies of traffic and transportation theory: a brief review, Phys. Stat. Mech. Appl. 392 (2013) 5261–5282, http://dx.doi.org/10.1016/j.physa.2013.06.004.
[42] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2006.
[43] B. Hoh, M. Gruteser, R. Herring, J. Ban, D. Work, J.-C. Herrera, A.M. Bayen, M. Annavaram, Q. Jacobson, Virtual Trip Lines for Distributed Privacy-preserving Traffic Monitoring, ACM, 2008, http://dx.doi.org/10.1145/1378600.1378604, pp. 15–28.
[44] R. Li, G. Rose, M. Sarvi, Evaluation of speed-based travel time estimation models, J. Transp. Eng. 132 (2006) 540–547, http://dx.doi.org/10.1061/(ASCE)0733-947X(2006)132:7(540).

[45] C. Cortés, R. Lavanya, J.-S. Oh, R. Jayakrishnan, General-Purpose Methodology for Estimating Link Travel Time with Multiple-Point Detection of Traffic, Transp. Res. Rec.: J. Transp. Res. Board 1802 (2002) 181–189, http://dx.doi.org/10.3141/1802-20.
[46] J. van Lint, N. van der Zijpp, Improving a travel-time estimation algorithm by using dual loop detectors, Transp. Res. Rec.: J. Transp. Res. Board 1855 (2003) 41–48, http://dx.doi.org/10.3141/1855-05.
[47] D. Ni, H. Wang, Trajectory reconstruction for travel time estimation, J. Intell. Transp. Syst. Technol. Plan. Oper. 12 (2008) 113–125, http://dx.doi.org/10.1080/15472450802262307.
[48] C. Shi, B.Y. Chen, Q. Li, Estimation of travel time distributions in urban road networks using low-frequency floating Car data, ISPRS Int. J. Geoinf. 6 (2017) 253, http://dx.doi.org/10.3390/ijgi6080253.