Freeway travel time estimation based on the general motors model: a genetic algorithm calibration framework

Shu Yang[1], Jishun Ou[2] ✉, Yiheng Feng[3], Yuan Wang[4]

[1]*School of Transportation, Southeast University, No.2, Southeast University Road, Nanjing 211189, People's Republic of China*
[2]*Intelligent Transportation System Research Center, Southeast University, No.2, Southeast University Road, Nanjing 211189, People's Republic of China*
[3]*Transportation Research Institute, University of Michigan, 2901 Baxter Rd, Ann Arbor, Michigan 48109, USA*
[4]*Department of Civil & Environmental Engineering, University of South Florida, 4202 E Fowler Avenue, ENB 118, Tampa, Florida 33620, USA*
✉ *E-mail: jishun@seu.edu.cn*

Abstract: Travel time estimation plays an important role in freeway performance assessment and reliability management. Conventional estimation methods based on speed information do not consider the level of congestion during modelling, leading to unreliable estimations under congested conditions. This study presents a freeway travel time estimation method based on the general motors (GM) model, where the concept 'virtual vehicle' was proposed to apply macroscopic traffic flow data to the microscopic GM model. First, a GM-based travel time estimation (GMTTE) model was developed to estimate link travel time. Three sets of parameters l, m and α were separately used under free-flowing, congestion and transition conditions. Next, two corridor estimation models, namely GMTTE time-slice-based (GMTTE-TSB) model and GMTTE continuous-speed (GMTTE-CS) model, were proposed. To calibrate the three parameters, a genetic algorithm framework was developed. The proposed method was calibrated and evaluated using traffic data collected from seven freeway segments in the greater St. Louis, Missouri region. Results show that (i) the GMTTE-CS model outperformed the GMTTE-TSB model; (ii) the optimal parameters of the GMTTE-CS model under free-flowing and congested conditions are ($l = 1.1$, $m = 2.0$, $\alpha = 8.0$) and ($l = 1$, $m = 0.1$, $\alpha = 8.0$), respectively; (iii) the GMTTE-CS model outperformed the instantaneous and time-slice models under congested conditions, while showing similar accuracies with the two models under free-flowing and status-transition conditions; and (iv) the optimal time interval was found to be 9–10 min for free-flowing conditions and 6–7 min for congested conditions.

1 Introduction

Travel time is a key measure for freeway performance assessment and reliability management. Local agencies are often required to report travel time information to higher-level transportation agencies [1, 2]. Travel time information can also help the public make reasonable travel choices, and serves as one of the fundamental indicators in transportation planning and facility accessibility evaluation [3, 4]. Accurate and reliable travel time estimation benefits both the public and transportation agencies.

Many manual strategies have been developed to directly collect freeway travel times, including GPS-based probe vehicles, test vehicles, vehicle signature matching and so on [5]. Besides, the use of Intelligent Transportation System (ITS) data to indirectly estimate travel time is becoming more popular, owing to two major facts. First, most state-level departments of transportation have installed ITS sensors on major freeways. The nationwide network of the ITS sensors (e.g. microwave radar sensors and passive infrared sensors) ensures the availability of a large amount of ITS data. Second, numerous mathematical models have been presented to quantify freeway traffic conditions. These models relieve traffic engineers and practitioners from labour-intensive manual travel time collection work.

The most fundamental definition for travel time on a freeway link is travel time equals the link length divided by speed. Various models define the speed in different ways. For example, the Instantaneous model considers the speed as the average of upstream and downstream speeds at vehicle departure times. The time-slice model improves the instantaneous model by using the travel time on the ith link to adjust the departure time for the $(i + 1)$th link. The calculated speed then becomes the average of upstream and downstream speeds at the adjusted departure times. The dynamic time-slice model [6] is an advanced version of the time-slice model that recursively estimates the travel time on the

ith link and adjusts the departure time for the $(i + 1)$th link. A common assumption of the three models is that speed is constant within a link. As this is not always the case, Van Lint and Zijpp [7] refined the above models by using a linear transformation to estimate the changing speeds within links. Also, the piecewise truncated quadratic functions were employed by Sun *et al.* [8] to calculate speeds within freeway links. Similarly, Ni and Wang [9] constructed speed surfaces to reconstruct vehicle trajectories for travel time estimation. The models mentioned above are easy to understand and popular in practical applications, and can be deemed as trajectory-based travel time estimation models [10].

Some other categories of methods based on speed information were also explored by researchers, including vehicle reidentification methods, traffic-flow-theory based methods and data-driven based methods. Most of the vehicle reidentification methods require the time-consuming examinations of electronic signatures collected from ITS sensors, while the data-driven based methods need a large amount of high-quality data for model training and validation [10]. The traffic-flow-theory based methods are built on classical traffic flow models. For instance, Coifman [11] used a two-regime traffic flow model to reconstruct vehicle trajectories, and further inferred freeway travel times. Numerical results show the estimated travel times are consistent with the ground truth travel times under uninterrupted traffic conditions. However, the model fails to deal with the cases where a queue partially covers a link. The traffic-flow-theory based methods have been shown the advantages in terms of estimation accuracy, especially under congested conditions.

Gazis *et al.* [12] and Herman *et al.* [13] proposed a microscopic model called the general motors (GM) model to mathematically describe the responses of the following vehicles to leading vehicles. Meanwhile, the relationship between the GM model and macroscopic traffic flow models has also been established [12–14]. For instance, the well-known Greenshield's and Greenberg's traffic
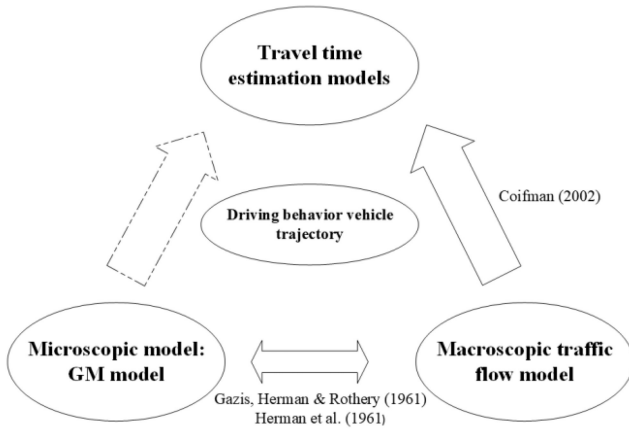
Fig. 1 *Relationships among microscopic, macroscopic and travel time estimation models*

flow models can be derived from the GM model. The study by Coifman [11] shows that travel time estimation models can be built based on macroscopic traffic flow models. Fig. 1 depicts the relationship between the microscopic model and macroscopic traffic flow models, as well as the relationship between the macroscopic traffic flow models and travel time estimation models.

Li *et al.* [15] compared the estimation results of the instantaneous, time-slice, dynamic time-slice and linear transformation models [7] with the ground truth travel times, concluding that 'the level of error declines with increasing level of congestion'. The conclusion demonstrates the accuracy of travel time estimation methods is closely associated with the level of congestion. In other words, the performance of the estimation methods can be further improved if the level of congestion is properly considered during modelling.

Inspired by the studies by Coifman [11] and Li *et al.* [15], this paper presents a freeway travel time estimation method based on the GM car-following model. Unlike the conventional models based on speed information, e.g. instantaneous model, time-slice model and dynamic time-slice model, the proposed method considers the level of congestion by incorporating three representative parameters, i.e. distance headway exponent *l*, speed exponent *m*, and sensitivity coefficients *α*, into the modelling process. The combined use of the three parameters can characterise and reflect various traffic conditions on freeways, and can thus be utilised to improve travel time estimation models. To the best knowledge of the authors, little research has been conducted to estimate travel time using the GM model. Therefore, the travel time estimation method proposed in the study is developed by the GM model to complete the relationships shown in Fig. 1. A previous study conducted by Yang *et al.* [16] demonstrates the feasibility and effectiveness of the GM-based method. However, the model parameters were empirically determined. In this study, we focus on (i) proposing innovative freeway travel time estimation models based on a car-following model with three parameters built-in; (ii) extending the previously proposed method by using a different speed initialisation strategy for the start of freeway links; (iii) proposing an evolutionary algorithm to calibrate the three parameters in the proposed travel time estimation model; and (iv) determining an optimal aggregation time interval for freeway travel time estimation.

To avoid the ambiguity regarding the concept of travel time estimation, three notes are listed below:

- No travel time prediction efforts are being made in this study. Travel time estimation and travel time prediction were sometimes interchangeably used in previous studies [10]. This study aims to use past speed data to estimate travel time.
- Similar to the instantaneous and time-slice models, the proposed models intend to estimate freeway travel time as a general indicator of traffic conditions, e.g. level of service, rather than evaluating individual trip travel time [10]. The estimated travel

times are commonly used by Advanced Traveller Information Systems (ATISs) and Advanced Traffic Management Systems (ATMSs) and displayed by Dynamic Message Signs (DMSs) on freeway roadsides.
- Although car-following models are widely used for traffic simulation, they were originally presented for modelling the driving behaviours of leader and follower vehicles. Therefore, no simulation steps are required by the proposed travel time estimation method. Meanwhile, the selected car-following model (i.e. the GM model) is a deterministic car-following model, meaning the results of travel time estimation would not be changed after the three parameters are determined.

The remaining of the paper is organised as follows. Section 2 provides the methodology of the proposed method, where the concept of 'virtual vehicle' is presented and the link and corridor travel time estimation models based on the GM model are developed. A genetic algorithm (GA) calibration framework is proposed to optimally determine the parameters of the proposed models. In Section 3, the study freeway corridors and collected data are described, while the experimental studies in Section 4 are conducted to evaluate the proposed method. Conclusions and future work remain in Section 5.

2 Methodology

2.1 Link travel time estimation based on the GM model

The GM model is a microscope car-following model, which is typically used to measure kinetic responses between two consecutive travelling vehicles on roadways. Three parameters, i.e. distance headway exponent *l*, speed exponent *m*, and sensitivity coefficients *α*, were defined to characterise car-following behaviours [12, 13]. Equations (1)–(3) describe the mathematical form of the GM model.

$$v_n^t = v_n^{t-\Delta T} + a_n^{t-\Delta T} * \Delta T, \tag{1}$$

$$x_n^t = x_n^{t-\Delta T} + v_n^{t-\Delta T} * \Delta T + \frac{1}{2}a_n^{t-\Delta T}\Delta T^2, \tag{2}$$

$$a_{n+1}^t = \left[\frac{\alpha_{l,m}\left(v_{n+1}^t\right)^m}{\left(x_n^{t-\Delta T} - x_{n+1}^{t-\Delta T}\right)^l} \right] * \left(v_n^{t-\Delta T} - v_{n+1}^{t-\Delta T}\right), \tag{3}$$

where v_n^t is the instantaneous speed of the *n*th vehicle at time *t*, a_n^t is the instantaneous acceleration rate of the *n*th vehicle at time *t*, ΔT is the time interval, x_n^t is the travelling distance of the *n*th vehicle at time *t*, *l* is the distance headway exponent between [−1, 4], *m* is the speed exponent between [−2, 2], and $\alpha_{l,m}$ is the sensitivity coefficient.

Equations (1) and (2) are based on classical kinematic equations and used to estimate the *n*th vehicle's (leading vehicle) instantaneous speed and travelled distance. One important kinetic response measure is the acceleration rate. Gazis *et al.* [12] and Herman *et al.* [13] summarised previous studies and proposed the generalised form (3) to estimate the acceleration rate of the (*n* + 1)th vehicle (following vehicle). The generalised form is associated with the current kinetic statuses of both of the leading and following vehicles. By using (1)–(3) iteratively, vehicle kinetic statuses, as well as travel distance and time, can be calculated simultaneously. However, the inputs of the GM model require detailed vehicle movement data, e.g. instantaneous speed and acceleration rate at each time interval. To establish the GM model using the ITS traffic sensor data which is usually aggregated over a relatively long duration (typically 20 or 30 s), the concept of 'virtual vehicle' is proposed.

Fig. 2 shows a freeway link bounded by upstream and downstream ITS sensors S_u and S_d. Assume two virtual vehicles, a virtual leading (VL) vehicle and a virtual following (VF) vehicle, travel on the freeway link. The initial headway of the two virtual vehicles at the starting time t_0 is set as the link length. The speeds
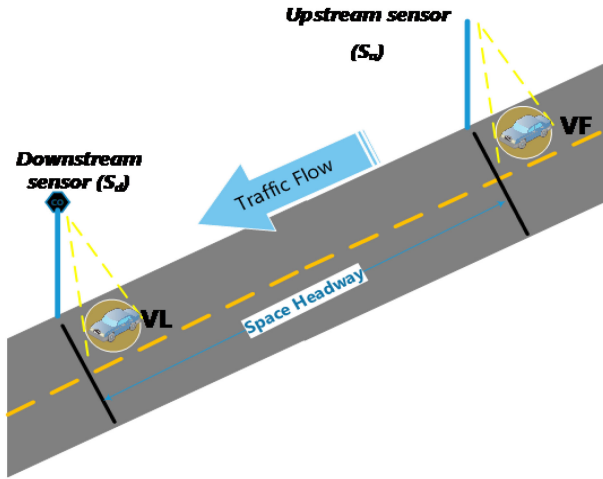
Fig. 2 *Virtual vehicle illustration*

of the VL and VF vehicles at t_0 are initialised as the downstream and upstream speeds collected at the time t_0, denoted as $v_{S_d}^{t_0}$ and $v_{S_u}^{t_0}$, respectively. After t_0, the speed of the VL vehicle is still determined by the downstream sensor S_d, whereas the speed of the VF vehicle is calculated based on the GM model in response to the kinetic changes of the VL vehicle. As mentioned above, the roles of the upstream sensors S_u are (i) to contribute to the initial speed of VF vehicles and (ii) to serve as a location/position reference for VF vehicles. Based on the above iterative process, the time–space diagram of the VF vehicle can be generated and, accordingly, its link travel time can be derived. The link travel time is finally determined as the VF vehicle's link travel time. The link travel time is defined as the travel time between two consecutive points on freeways. In the study, the consecutive points are selected to be the locations of traffic sensors. Note that the aforementioned kinetic mechanism of the virtual vehicles could be just one of possible ways to build the microscopic car-following models with aggregated ITS data. The other mechanisms can also be introduced, but will not discuss in this study.

It should be indicated that traffic sensors usually report aggregated speed at a certain time interval (e.g. 20 s, 30 s or 1 min). In this study, the used speed data is aggregated every 30 s. An interpolation algorithm is required to estimate the instantaneous speed of the VL vehicle at the time t. Equation (4) shows the mathematical form of the adopted interpolation algorithm

$$v_{VL}^t = \frac{v_{S_d}^{T_{i+1}} - v_{S_d}^{T_i}}{T_{i+1} - T_i} * (t - T_i) + v_{S_d}^{T_i} \ (T_i \leq t \leq T_{i+1}), \tag{4}$$

where v_{VL}^t is the instantaneous speed of the VL vehicle at the time t, T_i is the timestamp of the collected ITS sensor data, and $v_{S_d}^{T_i}$ is the speed collected from S_d at the time T_i.

Based on (1)–(4) and the concept of 'virtual vehicle', a freeway link travel time estimation model, namely GM-based travel time estimating (GMTTE) model, is proposed. The mathematical form of the GMTTE model can be simply denoted as

$$[\text{TT}_i, \ v_{\text{TT}_i}] = \text{GMTTE}(l, \ m, \ \alpha, \ v_{S_d}, \ v_{S_u}), \tag{5}$$

where l, m, and α are the parameters to be calibrated, and v_{S_d} and v_{S_u} are the input aggregated speed data at downstream and upstream ITS sensors, respectively. The outputs of (5) consist of the travel time on the ith link, TT_i, and the instantaneous speed of the VF vehicle at the end of the ith link, v_{TT_i}.

2.2 Corridor travel time estimation

Two corridor travel time estimation models based on the GMTTE model, are proposed in the following two subsections. Assume a

corridor is composed of $(n + 1)$ links, denoted as ith link, $(i + 1)$th link, $(i + 2)$th link, …, $(i + n)$th link, respectively. The corridor travel time is defined as the sum of the travel time on consecutive freeway links. The first model is similar to the time-slice model (which improves the instantaneous model by using the travel time on the ith link to adjust the departure time for the $(i + 1)$th link), named GMTTE time-slice-based (GMTTE-TSB) model. The model resets the speed at the beginning of the ith link to be the speed of downstream sensors at entering the ith link. The second model is different from the first model. The speed of the VF vehicle at the end of ith link is used as the initial speed at the beginning of the $(i + 1)$th link. Therefore, the speed profile of the VF vehicle is continuous. The second model is then named GMTTE continuous-speed (GMTTE-CS). The main distinction between the two models lies in the different initial speeds of the VF vehicles at the beginning of each link. The GMTTE-TSB model uses the time-slice model's strategy to initialise the starting speed, while the GMTTE-CS model uses the speed at the end of the ith link, v_{TT_i}, to initialise the starting speed at the beginning of the $(i + 1)$th link.

2.2.1 GMTTE-TSB model: The GMTTE-TSB model uses the following steps to estimate corridor travel time:

Step 1: Initialise $i = 0$.
Step 2: Use the link model $\text{GMTTE}(l, \ m, \ \alpha, \ v_{S_{d_i}}, \ v_{S_{u_i}})$ to estimate the travel time on the ith link, TT_i.
Step 3: Adjust the departure time for the $(i + 1)$th link based on TT_i.
Step 4: Estimate the travel time TT_{i+1} using the speed data at the adjusted departure time.
Step 5: Let $i = i + 1$ and repeat Steps 2–4 to estimate travel time on each link.
Step 6: Sum all link travel times to obtain the corridor travel time.

2.2.2 GMTTE-CS model: The following five steps are included in the GMTTE-CS model to estimate corridor travel time:

Step 1: Initialise $i = 0$.
Step 2: Use the link model $\text{GMTTE}(l, \ m, \ \alpha, \ v_{S_{d_i}}, \ v_{S_{u_i}})$ to estimate TT_i and v_{TT_i}.
Step 3: Initialise the starting speed on the $(i + 1)$th link with v_{TT_i}.
Step 4: Use $\text{GMTTE}(l, \ m, \ \alpha, \ v_{S_{d_{i+1}}}, \ v_{\text{TT}_i})$ to estimate the $(i + 1)$th link travel time.
Step 5: Let $i = i + 1$ and repeat Steps 2–4 to estimate travel time on each link.
Step 6: Sum all link travel times to obtain the corridor travel time.

2.3 GA calibration framework

To determine the optimal parameter set (l, m, α) for the proposed models, a GA calibration framework is proposed. Two performance measures, i.e. mean absolute error (MAE) and mean absolute percentage error (MAPE), are used to design the fitness function of the GA.

2.3.1 Performance measures: The MAE and MAPE measures are defined as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|, \tag{6}$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \tag{7}$$

where y_i is the true value of the ith data sample, $\hat{y}_i$ is the estimated value of the ith data sample, and n is the number of data samples.
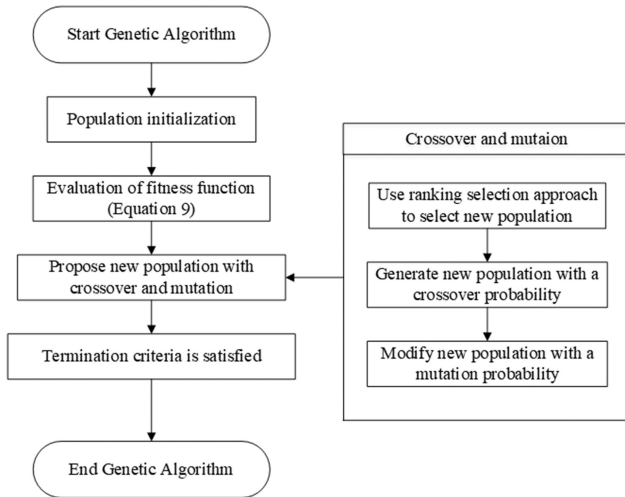
Fig. 3 *Flowchart of designed GA*

2.3.2 GA design: Because the ITS sensor data was reported every 30 s and the ground truth travel times were collected at vehicle level (2–5 individual vehicle travel times were collected within a minute), the estimated and ground truth travel times were both aggregated into a certain time interval before calibrating the proposed models. Aggregation over the same time interval enable the calculation of MAEs and MAPEs and hence the evaluation the proposed estimation models. The objective of the GA is to solve the following optimisation problem:

$$f(l, m, \alpha) = \arg\min_{l, m, \alpha} MA_j(l, m, \alpha), \qquad (8)$$

where MA_j represents the model accuracy associated with the time interval j, and could be either MAE_j or $MAPE_j$ in this study.

To solve the above problem effectively, a proper fitness function is required by the GA, so that new generations can evolve towards the best solution. As potential fitness functions could include MA_j with any value of j, a reasonable guess would be that the selection of j can affect the calibration results. For example, a set of (l, m, α) can minimise a fitness function when MA_5 is used, whereas it may not be able to minimise a fitness function with a 10-min aggregation interval. Furthermore, another reasonable guess may be raised: the value of j could affect the accuracy of the aggregated travel time due to the trade-off between traffic dynamics and travel time estimation noises. For instance, if j is set to be a small value (e.g. 1 or 2 min), the aggregated travel time would vary significantly and noises might be introduced; on the other hand, if j is set as a relatively large value (e.g. 15 or 20 min), the aggregated travel time would vary less and fail to capture travel time dynamics. Therefore, proper j needs to be determined carefully. However, a few guidelines are available to assist in selecting the optimal j. Given this, a fitness function with weighted accuracies is introduced, as shown in the following equation:

$$\text{Fitness}(l, m, \alpha) = \frac{\sum_{j=2}^{K} W_j * MA_j}{\sum_{j=2}^{K} W_j}, \qquad (9)$$

where W_j is the weight of MA_j, and K is the upper bound of j. The lower bound is selected as 2 because the ground truth travel times are difficult to collect within time intervals less than 2 min.

Fig. 3 depicts the proposed GA calibration framework. As the figure shows, an initial population is composed of a set of chromosomes (solutions) and produced by generating random numbers within the ranges of the parameters to be solved. Starting with the initial population, the GA carries out a process of fitness-based selection, crossover and mutation, to produce a successor population. The selection operation is to select a group of candidate solutions to breed a new generation while the crossover operation produces offspring by exchanging the selected chromosomes among genes (a gene is a unit code of a chromosome). The

mutation operation changes the value of some chromosomes at random and allows the GA to jump out of a local optimum by occasionally exploring other regions. As the above process is iterated, a sequence of generations evolves, and the average fitness of the chromosomes tends to increase until the termination criterion is reached. With such a searching strategy, the GA is capable of evolving the best solution for the calibrated models. In the study, the crossover probability and mutation probability were set as 0.1 and 0.01, respectively, and the ranking selection mechanism [17] was adopted to select the prospective offspring.

2.3.3 Weight updating using Metropolis–Hastings algorithm: The parameter W_j defined in (9) follows a probability distribution on a time interval i, denoted as $p(i) = W_i$. Various traffic conditions could result in different optimal j and weights. To this end, two prior distributions of j are used to characterise different traffic conditions.

For the free-flowing conditions, a uniform distribution is adopted, assuming each j could be optimal and its associated MA_j is evenly contributed to the fitness function. The distribution is formulated as follows:

$$p(j) = \begin{cases} \dfrac{1}{\sigma_2 - \sigma_1} & \sigma_1 \leq j \leq \sigma_2 \\ 0 & j < \sigma_1 \text{ or } j > \sigma_2 \end{cases}, \qquad (10)$$

where σ_1 and σ_2 are 2 and 15 min, respectively. Any j greater than 15 min is considered inadequate to capture travel time dynamics.

For the congested conditions, a log-normal distribution is used as the prior distribution, taking the form

$$p(j) = \frac{1}{j\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln j - \mu)^2}{2\sigma_l^2}\right), \qquad (11)$$

where $\mu = 3$ and $\sigma_l = 1$ in the study. The distribution indicates that a smaller j is more likely to be the optimal j and thus given more weights.

It is possible that the two prior distributions may not thoroughly represent the distributions on j, an updating mechanism is introduced to update the distributions by incorporating the best i. The best i is set as the time interval with the lowest MA_j, and considered as the optimal candidate j_{can}. Each population in the GA produces an optimal candidate. Let j_{can}^M be the optimal candidate in the Mth iteration. The updating mechanism allows the fitness function to evolve by incorporating j_{can}^M after the Mth iteration. Updating the two prior distributions with j_{can}^M provides two posterior distributions for the optimal j. The maximum a posteriori (MAP) method and Metropolis–Hasting algorithm [18] have been widely applied to update prior distributions with new observations. However, the MAP method is limited to the prior distribution selection and requires the used prior and posteriori distributions are conjugated. Since the uniform and log-normal distributions have no corresponding conjugate distributions and no closed mathematical expressions can be derived to update the distributions, the Metropolis–Hasting algorithm is employed in this study. The algorithm samples from unknown distributions with a ratio of likelihoods r. The following two steps describe the calculation process of a posterior distribution.

Step 1: Propose a new j. The probabilistic relationship between the known optimal j_{old} and newly proposed optimal j_{new} is described with a Gaussian distribution $N(j_{\text{old}}, \sigma_r)$. This proposal step is known as a random walk process. If j_{new} is accepted, j_{old} is set to be j_{new}; otherwise, j_{old} remains unchanged. The initial j_{old} can be any non-negative number and was set as 10 in the study. Equation (12) shows the established relationship between j_{old} and j_{new}:

$$p(j_{\text{new}} | j_{\text{old}}) = N(j_{\text{old}}, \sigma_r), \qquad (12)$$

$$r = \frac{p(j_{\text{new}})\text{Likelihood}(j_{\text{can}} \mid j_{\text{new}})}{p(j_{\text{old}})\text{Likelihood}(j_{\text{can}} \mid j_{\text{old}})} \quad (13)$$

$$r_{\text{uniform}} = \frac{\dfrac{1}{\sigma_2 - \sigma_1} * \prod_m^M N(j_{\text{can}}^m \mid j_{\text{new}}, \ \sigma_a)}{\dfrac{1}{\sigma_2 - \sigma_1} * \prod_m^M N(j_{\text{can}}^m \mid j_{\text{old}}, \ \sigma_a)} \quad (14)$$

(see (15)) , where σ_r is the walk size, and was set as 5 in the study.

Step 2: Accept or reject the proposed optimal j based on $j_{\text{can}}^1, j_{\text{can}}^2, \ldots, j_{\text{can}}^M$. A ratio of likelihoods r is used to determine acceptance or rejection, as shown in (13). The Likelihood$(j_{\text{can}} \mid j_{\text{new}})$ is defined as the multiplication of the Gaussian distributions $N(j_{\text{can}} \mid j_{\text{new}}, \ \sigma_a)$. Equations (14) and (15) show the derived r using the two prior distributions defined in (10) and (11). If r is greater than 1, or r is greater than a random number between [0, 1] when r is less than 1, the proposed optimal j_{new} is accepted.

After repeating the two steps N times (2000 in the study), the two posterior distributions are updated after the Mth iteration. Since the two posterior distributions of the time interval j are obtained, the probability at j (i.e. the weight of the time interval j) can be calculated. Then, the newly calculated weight of j (W_j) replaces the old weight and cycles into the next iteration until convergence. It is worth noting that σ_r in (12) only affects the acceptance rate. The posterior distributions, as well as the corresponding statistics (e.g. mean, mode and standard deviation), are not affected. Similarly, σ_a in (14) and (15) only affects the acceptance rate as well. Fig. 4 illustrates the weight updating process.

3 Data collection and description

The traffic data collected from seven freeway segments in the greater St. Louis, Missouri region was used for both model calibration and evaluation. Two freeway corridors located on I-270 were selected for model calibration. Corridor 1 is an 11.59 km (7.2 miles) segment of I-270 Southbound, which has little recurrent congestion on Fridays. Given this, Corridor 1 on Friday mornings was used for a free-flowing scenario. Corridor 2 is a 5.95 km (3.7 miles) segment of I-270 Northbound, which consistently suffers from traffic congestion on weekdays. Accordingly, Corridor 2 on Tuesday mornings was used for a congested scenario. Two other freeway corridors, namely Corridor 3 and Corridor 4, were selected for the model evaluation of the free-flowing scenarios. No efforts have been made to collect ground truth data because travellers and traffic engineers are more interested in acquiring travel time information under congestion scenarios. However, relative comparisons for models will be performed under free-flowing scenarios. These two corridors were located on different freeway segments. Corridor 3 is a 4.35 km (2.7 miles) segment of I-270 Southbound, while Corridor 4 is a 5.79 km (3.6 miles) segment of I-64 Westbound. Also, two more freeway corridors, Corridors 5 and 6, were selected for evaluating the congested scenarios. The 'stop-and-go' traffic conditions were often observed on entire Corridors 5 and 6. The last corridor, Corridor 7, is an 8.85 km (5.5 miles) segment (4 links) of I-270 Southbound. Traffic congestions were observed on the first three links but mitigated on the fourth link. Corridor 7 was thus selected to evaluate the models when the traffic goes through the transition state.

Table 1 shows the detailed information of the seven corridors. Two types of data were collected, including ITS sensor data and actual travel times. The ITS sensor data was used to estimate travel time using the proposed models. In addition to the ITS sensors, surveillance cameras were installed on I-270. The vehicle matching
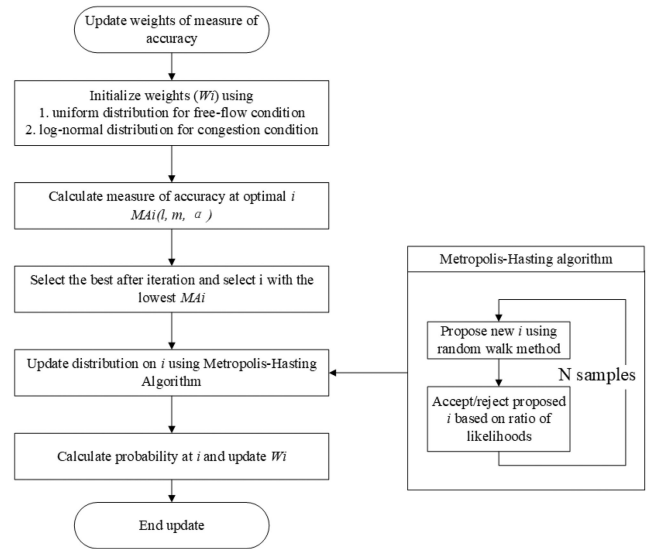

Fig. 4 *Flowchart of weight updating using Metropolis–Hastings algorithm*

approach proposed by Schroeder *et al.* [5] was implemented to collect the ground truth travel times. Note that the ground truth travel times are not available for the two free-flowing verification scenarios (i.e. Corridors 3 and 4). Similar to the work by Ni and Wang [9], comparative analyses with other estimation models were also conducted.

4 Model calibration and evaluation

4.1 Model calibration

The GA framework presented in Section 2.3 was carried out to calibrate the proposed GMTTE-TSB and GMTTE-CS models based on the MAE and MAPE measures. As each model was calibrated by using two measures on two study corridors, a total of eight calibration scenarios can be created, including:

Scenario 1a: GMTTE-TSB model using MAPE on Corridor 1.
Scenario 1b: GMTTE-TSB model using MAPE on Corridor 2.
Scenario 2a: GMTTE-TSB model using MAE on Corridor 1.
Scenario 2b: GMTTE-TSB model using MAE on Corridor 2.
Scenario 3a: GMTTE-CS model using MAPE on Corridor 1.
Scenario 3b: GMTTE-CS model using MAPE on Corridor 2.
Scenario 4a: GMTTE-CS model using MAE on Corridor 1.
Scenario 4b: GMTTE-CS model using MAE on Corridor 2.

Table 2 shows the optimal sets of (l, m, α) and the optimal j for Scenarios 1 and 2, while Table 3 shows the corresponding results for Scenarios 3 and 4. According to Brackstone and McDonald [14], α could be mathematically cancelled out for calibrating the GM model, and only the parameters l and m need to be considered. To examine whether the selection of α has significant impacts on travel time estimation, the effects of selection of α and the best fitness are also listed. Detailed findings are summarised below:

(i) The effects of selection of α ranging from 3 to 14 were tested. The results show α had minor effects on the accuracy of estimated travel times. For example, the MAEs in Scenario 1a range from 1.54–1.58 s and in Scenario 1b range from 11.57 to 11.69 s. Therefore, the effects of selection of α for all scenarios can be ignored. This conclusion is consistent with the previous study [14].

$$r_{\text{lognorm}} = \frac{\dfrac{1}{j_{\text{new}}\sqrt{2\pi}}\exp\left(-\dfrac{(ln\,j_{\text{new}} - 3)^2}{2}\right) * \prod_m^M N(j_{\text{can}}^m \mid j_{\text{new}}, \ \sigma_a)}{\dfrac{1}{j_{\text{old}}\sqrt{2\pi}}\exp\left(-\dfrac{(ln\,j_{\text{old}} - 3)^2}{2}\right) * \prod_m^M N(j_{\text{can}}^m \mid j_{\text{old}}, \ \sigma_a)} \quad (15)$$

(ii) Scenarios 1b and 2b have similar results regarding the selection of an optimal set ($l = 0.5$, $m = -2.0$, $\alpha = 8.0$) for the GMTTE-TSB model. Although Scenarios 3b and 4b generated two different optimised sets ($l = 1$, $m = 0.1$, $\alpha = 8.0$) and ($l = 1$, $m = 0.1$, $\alpha = 6.0$), respectively. However, the only difference is the selection of α. To keep the following validation steps consistent and easy, the optimal set ($l = 1$, $m = 0.1$, $\alpha = 8.0$) was used for the GMTTE-CS model because of the minor effect of α.

(iii) The optimal sets of parameters under free-flowing traffic conditions vary. It was found that the four optimal sets of (l, m, α) in Tables 2 and 3 produced low errors. Essentially, any of the four optimal sets under free-flow conditions can be interchangeably used to produce travel time with low errors. Therefore, the parameter set ($l = 1.1$, $m = 2.0$, $\alpha = 8.0$) is used hereafter to verify the models.

The best fitness values in Tables 2 and 3 demonstrate the positive performance of the two calibrated models. To compare with other freeway travel time estimation models based on speed information, two conventional models, i.e. the instantaneous and time-slice models, were implemented for the same study corridors. The comparison results are shown in Fig. 5. Several findings are summarised below:

(i) The four models performed similarly under free-flowing conditions. The low MAPEs (<3%) and MAEs (<12 s) in Figs. 5a and b indicate that the performances of the four models are satisfactory, and all of the models can accurately estimate travel times under free-flowing conditions.

(ii) Estimating travel time under congested conditions is more challenging. Figs. 5c and d show the performances of the four models under congested conditions. The two conventional models have higher MAPEs (>20%) and MAEs (>90 s). In contrast, lower MAPEs and MAEs were achieved by the GMTTE-TSB and GMTTE-CS models. Moreover, it was observed that the GMTTE-CS model is superior to the GMTTE-TSB model. Therefore, the GMTTE-CS model with the optimal parameters

Table 1 Study corridors

Corridors	Time period	Number of ground truth travel times	Corridor descriptions
Data used for model calibration (training data)			
Corridor 1(three lanes)(free-flowing condition)	7–8 AM, 12 December 2014, Friday	183	direction: Southbound, length: 11.59 km (7.2 miles), Dorsett Road & I-270 – Clayton Road & I-270, six links
Corridor 2 (four lanes) (congested condition)	7:50–8:50 AM, 16 December 2014, Tuesday	285	direction: Northbound, length: 5.95 km (3.7 miles), Big Bend Road & I270 – Clayton Road & I-270, three links
Data used for model verification (test data)			
Corridor 3(three lanes)(free-flowing condition)	7–8 AM, 13 December 2014, Saturday	NA (relative comparison)	direction: Southbound, length: 4.35 km (2.7 miles), Dorsett Road & I270 – Olive Blvd & I270
Corridor 4(four lanes)(free-flowing condition)	11:30–13:30, 25 September 2014, Thursday	NA (relative comparison)	direction: Westbound, length: 5.79 km (3.6 miles), Sarah Street & I270 – McCausland Ave & I270
Corridor 5(four lanes) (congested condition)	7–8 AM, 17 December 2014, Friday	163	direction: Northbound, length: 8.85 km (5.5 miles), I44 & I270 – Clayton Road & I270, four links
Corridor 6(four lanes) (congested condition)	7–8 AM, 17 December 2014, Friday	123	direction: Northbound, length: 2.57 km (1.6 miles), I44 & I270 – Big Bend Road & I270), single link
Corridor 7(four lanes)(state transition condition)	15:00–15:50, 17 December 2014, Friday	72	direction: Southbound, length: 8.85 km (5.5 miles), Clayton Road & I270 – I44 & I270, four links

Table 2 Calibration results of the GMTTE-TSB model

		Corridor 1		Corridor 2	
		MAPE [Scenario 1a]	MAE [Scenario 2a]	MAPE [Scenario 1b]	MAE [Scenario 2b]
optimal j, min	mean	9.94	9.37	6.70	6.58
	median	9.95	9.38	6.70	6.58
optimal parameters	l	0.5	1.7	0.5	0.5
	m	0.7	2.0	−2.0	−2.0
	α	8.0	9.0	8.0	8.0
α effect ($3 \leq \alpha \leq 14$)	MAPE, %	[1.54, 1.58]	NA	[11.57, 11.69]	NA
	MAE (seconds)	NA	[6.42, 6.62]	NA	[51.36, 52.49]
best fitness	MAPE, %	1.54	NA	11.57	NA
	MAE, s	NA	6.42	NA	51.36

Table 3 Calibration results of the GMTTE-CS model

		Corridor 1		Corridor 2	
		MAPE [Scenario 3a]	MAE [Scenario 4a]	MAPE [Scenario 3b]	MAE [Scenario 4b]
optimal j, min	mean	9.21	9.38	7.35	6.99
	median	9.20	9.38	7.35	6.99
optimal parameters	l	1.1	1.7	1.0	1.0
	m	2.0	2.0	0.1	0.1
	α	8.0	5.0	8.0	6.0
α effect ($3 \leq \alpha \leq 14$)	MAPE, %	[1.27, 1.29]	NA	[6.86, 6.94]	NA
	MAE, s	NA	[5.94, 6.12]	NA	[31.51, 32.37]
best fitness	MAPE, %	1.27	NA	6.86	NA
	MAE, s	NA	5.94	NA	31.51
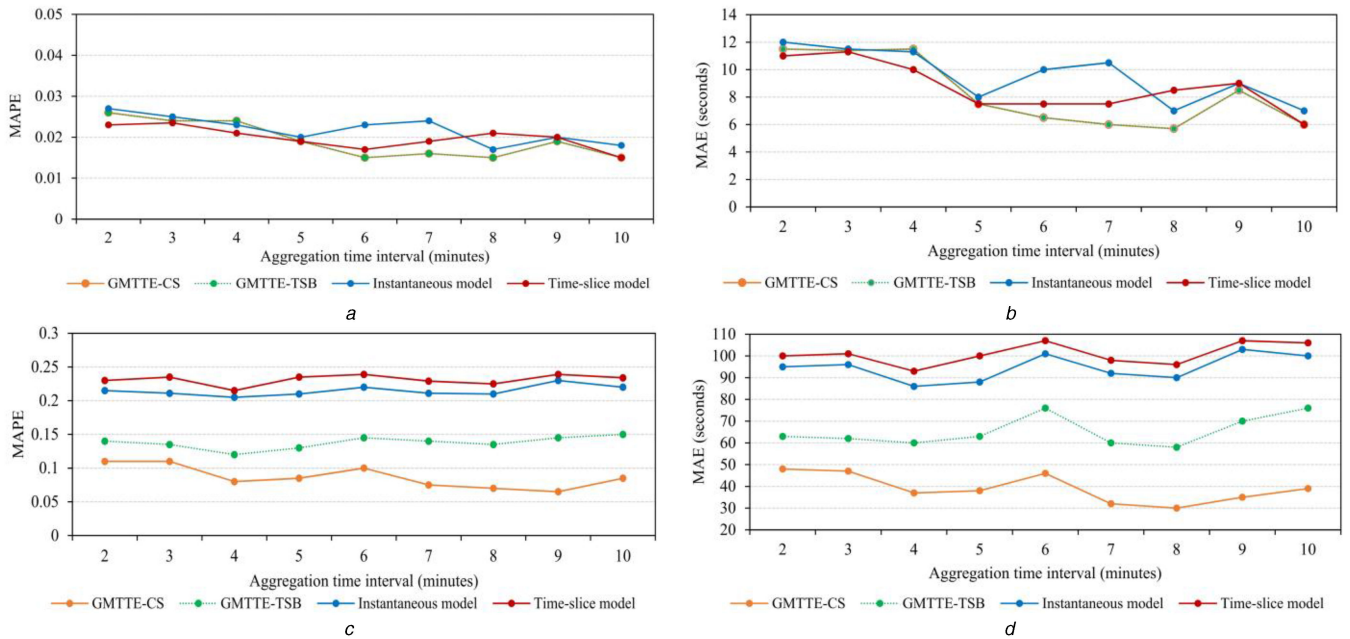
Fig. 5 *Model comparisons*

(a) MAPE comparisons on Corridor 1, *(b)* MAE comparisons on Corridor 1, *(c)* MAPE comparisons on Corridor 2, *(d)* MAE comparisons on Corridor 2

$(l = 1, \ m = 0.1, \ \alpha = 8.0)$ is used for model evaluation in later experiments.

4.2 Optimal aggregation time interval determination

To obtain consistent travel time estimates and capture traffic dynamic in travel time estimation, a proper aggregation time interval is required. Fig. 6 shows the posterior distributions based on the optimal j for the defined four scenarios. All of the posterior distributions are concentrated, and have small variances, though the probabilities of different time intervals are not equal. The higher the probability of a time interval, the more likelihood the time interval would be optimal j.

The uniform prior distribution indicates that any time interval has an identical likelihood to be optimal j under the free-flowing conditions. However, the corresponding posteriori distributions suggest that 9–10 min might be a good choice. For the congested conditions, it was suggested that 6–7 min could be selected as the optimal aggregation time interval.

4.3 Model evaluation

The GMTTE-CS model with $(l = 1.1, \ m = 2.0, \ \alpha = 8.0)$ was applied to the two free-flowing corridors (Corridors 3 and 4). Meanwhile, the model with $(l = 1, \ m = 0.1, \ \alpha = 8.0)$ was applied to the two congested corridors (Corridors 5 and 6) and Corridor 7 with status-transition conditions. Performance comparisons among the implemented models are shown in Fig. 7 and Table 4, respectively. The main findings are summarised as follows:

(i) As no ground truth travel time was available on Corridors 3 and 4, the corresponding quantitative comparisons were unavailable. However, based on Figs. 7a and b, minor differences can be observed on both corridors, implying the estimated models were similarly performed. This finding is consistent with the previous study by Li *et al.* [15].
(ii) The GMTTE-CS model outperformed the other two conventional models on both Corridors 5 and 6. The MAPEs of the two conventional models are ∼20%, while the MAPE is 7% lower using the GMTTE-CS model. Since Corridor 6 is a single-link scenario, the two conventional models performed the same, and their MAPEs are 14.1%. For the GMTTE-CS model, the MAPE is 7.97%. This demonstrates the performance advantage of the GMTTE-CS model over the conventional models under congested conditions.

(iii) The three models performed similarly on Corridor 7. They achieved higher accuracies, where all of the MAPEs are smaller than 4%.

Note that shorter links, meaning that more traffic sensors along a fixed freeway segment length, would capture more traffic dynamics, resulting in more accurate travel time estimates. Sparse deployment of traffic sensors may fail to capture traffic dynamics and result in less accurate travel times.

4.4 Statistical tests of model performance

In addition to the quantification of the differences between the estimation models using MAE and MAPE, statistical tests have also been conducted to statistically test the performance differences. Paired T-test is one of the common options to be applied. However, the T-test has a strong assumption, i.e. data samples should follow normal distributions. To relax this assumption, a non-parametric statistical test, namely the Wilcoxon signed-rank test, is applied in this study. The null and alternative hypotheses are H0: the difference between the paired travel time follows a symmetric distribution around zero; H1: the difference between the paired travel time does not follow a symmetric distribution around zero. The results of the statistical test are listed in Table 5. The results indicate that (i) no statistical difference of the model performances under free-flowing and transition conditions (i.e. Corridors 3, 4 and 7) is observed; and (ii) the statistical differences under congestion conditions (i.e. Corridors 5 and 6) are statistically significant. The combined results of MAE/MAPE and the statistical tests demonstrate that both GMTTE-TSB and GMTTE-CS outperformed the other two conventional freeway travel time estimation models under congested conditions.

5 Conclusion and future work

Freeway travel time estimation is an important topic in freeway traffic operation and management. The majority of previous studies developed trajectory-based, vehicle reidentification, traffic-flow-theory based and data-driven based methods to estimate freeway travel times. This study proposes a GM based freeway travel time estimation method with three parameters built-in, where two corridor travel time models are proposed by using different speed initialisation strategies on start links, namely GMTTE-TSB and GMTTE-CS models. Real-world data from seven corridors with various lengths (1.6–7.2 miles) and various traffic conditions (i.e. free-flow, congestion, and transition) were used for model
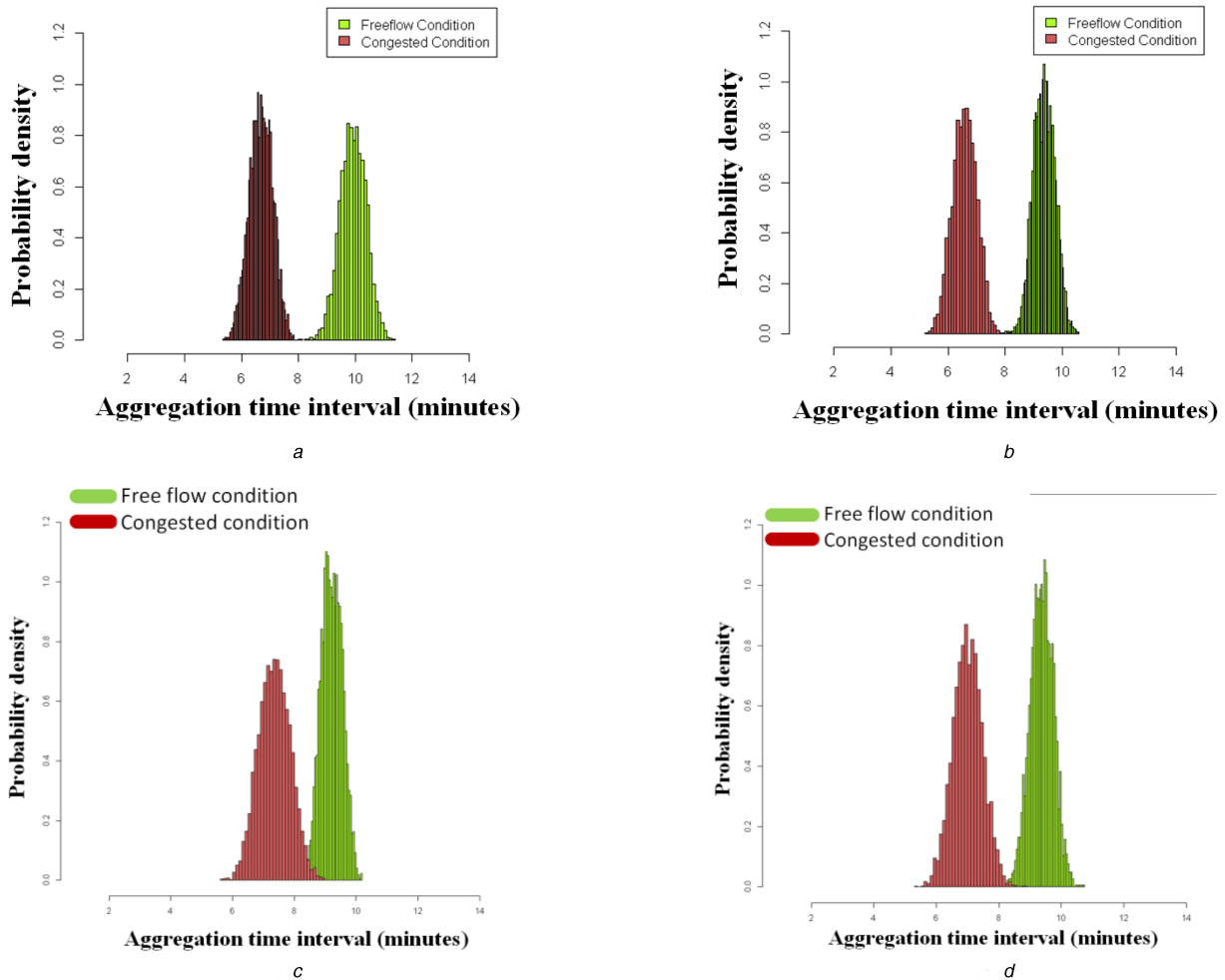
Fig. 6 *Distributions of optimal aggregation time interval*
(a) Scenarios 1a and 1b, *(b)* Scenarios 2a and 2b, *(c)* Scenarios 3a and 3b, *(d)* Scenarios 4a and 4b

calibration and verification. Among them, two corridors were used to calibrate the proposed models, while the rest of the five corridors were utilised to verify the calibrated models.

To obtain the optimal set of the three parameters inherited from the car-following model (i.e. the distance headway exponent (l), the speed exponent (m), and the sensitivity coefficient ($\alpha_{l,m}$)), a GA calibration framework was developed. The MAPE and MAE measures were used to design the fitness function of the GA. The calibration objective is defined to minimise estimation accuracy over a given aggregation time interval. Since different aggregation time intervals can lead to various optimal sets of parameters, a normalised weighted fitness function was designed to balance both travel time variances and dynamics. An update mechanism based on the Metropolis–Hastings algorithm was developed to update the weights of the fitness function.

The major contributions of this study are highlighted as follows:

(i) Two corridor travel time estimation models were proposed, i.e. GMTTE-TSB and GMTTE-CS, each of which has three parameters (l, m, α). Three optimised sets of the three parameters were used under different levels of congestion. Experimental analysis shows the selection of α has minor effects on travel time estimation, while the selection of l and m can significantly contribute to estimation accuracy. The calibration results suggest that the sets ($l = 1.1$, $m = 2.0$, $\alpha = 8.0$) and ($l = 1$, $m = 0.1$, $\alpha = 8.0$) are the optimal parameter sets for the proposed GMTTE-CS models.
(ii) Compared with the instantaneous and time-slice models, the proposed models show its performance advantage, especially for congested conditions. Moreover, the GMTTE-CS model is superior to the GMTTE-TSB model. The reasons could be summarised below:

- Speed may abruptly change within links using the GMTTE-TSB model. This abrupt change may lead to vehicle trajectory discontinuity.
- The vehicle trajectory generated from the GMTTE-CS model is closer to the actual vehicle movement because the estimated speed v_{TT} was continuously updated to ensure the speed is continuous.

(iii) A GA with an evolving fitness function was designed to optimise the three parameters and determine the best aggregation time interval. Few studies have explored and accounted for the effects of the optimal aggregation time interval. This study suggested that the optimal time intervals are 9–10 min for free-flowing conditions and 6–7 min for congested conditions. These time intervals could be further used for publishing traffic information (e.g. travel time or speed) through dynamic message signs. For example, agencies can predict travel times on a freeway segment over the next 9 or 10 min when the traffic is free-flowing or predict travel time over the next 6 or 7 min when traffic is congested.

The optimal set of parameters for the GMTTE-CS model was found to be ($l = 1$, $m = 0.1$, $\alpha = 8.0$). Meanwhile, Greenberg's macroscopic traffic flow model can be derived from the GM model when $l = 1$, $m = 0$ [14]. The minor difference of the parameters may inspire researchers to investigate the relationship between the GM and GMTTE model parameters. If the relationship can be mathematically established, many benefits would be obtained. For example, the GM model can be calibrated using aggregated ITS sensor data instead of detailed vehicle movement data. Future work could focus on investigating the relationship and discover the associated benefits. Furthermore, different kinetic mechanisms of virtual vehicles, traffic measures (e.g. volume and density) as well
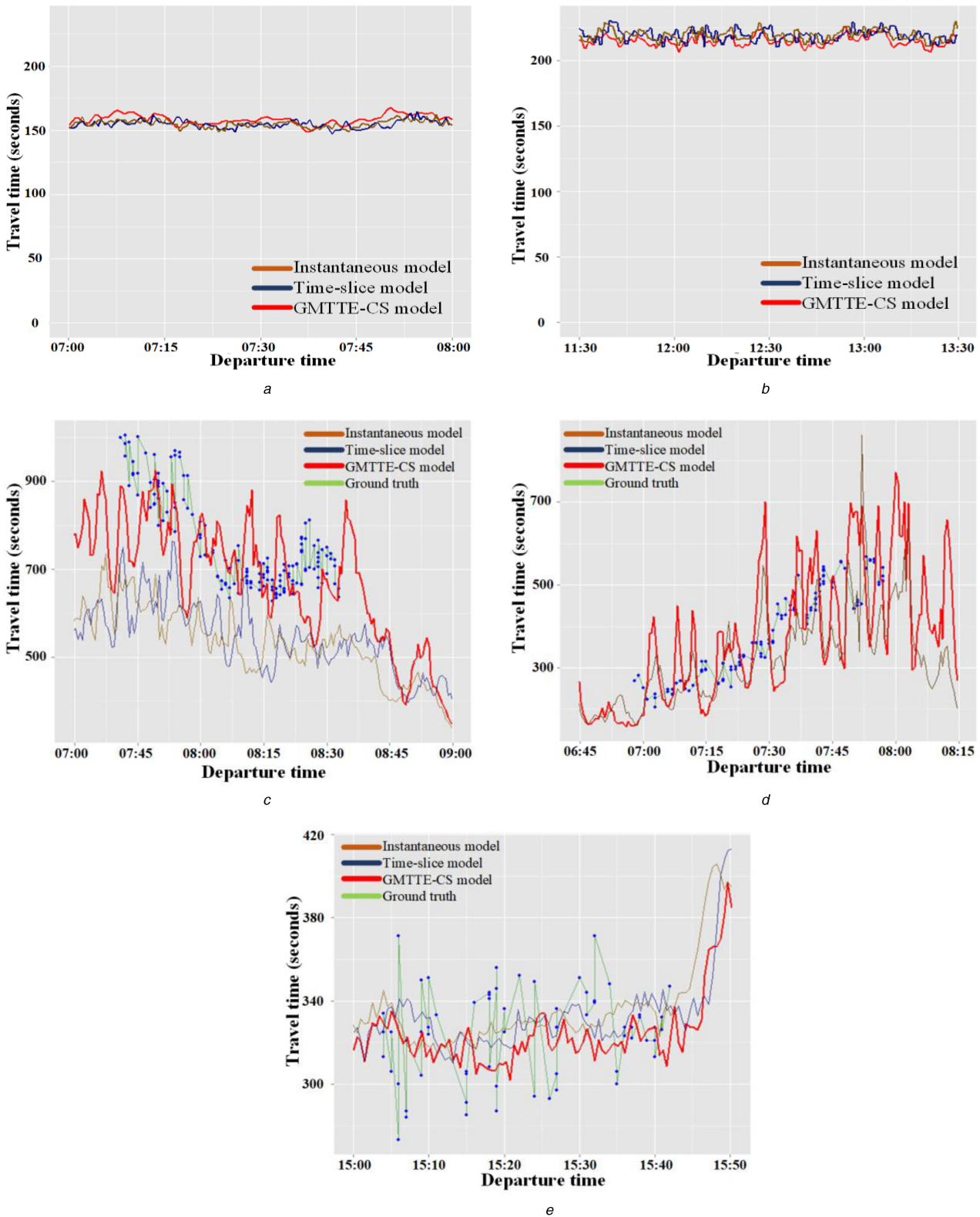
Fig. 7 *Evaluation results on five study corridors*
(a) Corridor 3, free-flowing condition, *(b)* Corridor 4, free-flowing condition, *(c)* Corridor 5, congested condition, *(d)* Corridor 6, congested condition, *(e)* Corridor 7, status transition condition

Table 4 Evaluation results of the GMTTE-CS, instantaneous and time-slice models

Study corridors	MAPE, %			MAE, s		
	GMTTE-CS	Instantaneous	Time-slice	GMTTE-CS	Instantaneous	Time-slice
Corridor 5	6.46	19.71	21.55	59.6	153.0	168.4
Corridor 6	7.97	14.1	14.1	25.2	58.9	58.9
Corridor 7	3.58	3.50	3.91	12.86	12.02	11.49

Table 5 Statistical tests results of the GMTTE-CS, instantaneous and time-slice models

Matrices	Comparison	Corridor (traffic condition)	p-values	Statistically significant difference
MAE	GMTTE-CS versus instantaneous model	Corridor 3 (free-flowing)	0.4881	no
		Corridor 4 (free-flowing)	0.9003	no
		Corridor 5 (congested)	3.2355×10^{-12}	yes
		Corridor 6 (congested)	8.4764×10^{-12}	yes
		Corridor 7 (transition)	0.3601	no
	GMTTE-CS versus time-slice model	Corridor 3 (free-flowing)	0.3708	no
		Corridor 4 (free-flowing)	1.0709	no
		Corridor 5 (congested)	6.5309×10^{-13}	yes
		Corridor 6 (congested)	2.6354×10^{-12}	yes
		Corridor 7 (transition)	0.2642	no
MAPE	GMTTE-TSB versus instantaneous model	Corridor 3 (free-flowing)	0.2193	no
		Corridor 4 (free-flowing)	0.58179	no
		Corridor 5 (congested)	4.2503×10^{-12}	yes
		Corridor 6 (congested)	5.2942×10^{-12}	yes
		Corridor 7 (transition)	0.1137	no
	GMTTE-TSB versus time-slice model	Corridor 3 (free-flowing)	0.3428	no
		Corridor 4 (free-flowing)	0.8474	no
		Corridor 5 (congested)	6.6534×10^{-12}	yes
		Corridor 6 (congested)	4.0669×10^{-12}	yes
		Corridor 7 (transition)	0.1924	no

as the usage of speed information from upstream sensors could be incorporated and tested to further improve the GM model. Finally, other car-following models can be explored to generalise the concept of 'virtual vehicle' for both rural and urban environments.

6 Acknowledgments

7 References

[1] Yang, S., Wu, Y.J.: 'Travel mode identification using bluetooth technology', *J. Intell. Transp. Syst.*, 2018, **22**, (5), pp. 407–421

[2] Yang, S., Chen, M., Wu, Y.J.*, et al.*: 'Machine learning approach to decomposing arterial travel time using a hidden Markov model with genetic algorithm', *J. Comput. Civ. Eng.*, 2018, **32**, (3), p. 04018005

[3] Geurs, K.T., Van Wee, B.: 'Accessibility evaluation of land-use and transport strategies: review and research directions', *J. Transp. Geogr.*, 2004, **12**, (2), pp. 127–140

[4] Páez, A., Scott, D.M., Morency, C.: 'Measuring accessibility: positive and normative implementations of various accessibility indicators', *J. Transp. Geogr.*, 2012, **25**, pp. 141–153

[5] Schroeder, B.J., Cunningham, C.M., Findley, D.J.*, et al.*: 'Manual of transportation engineering studies', 2010, ITE. ISBN-13: 978-1-933452-53-1; ISBN-10: 1-933452-53-6

[6] Cortés, C., Lavanya, R., Oh, J.S.*, et al.*: 'General-purpose methodology for estimating link travel time with multiple-point detection of traffic', *Transp. Res. Rec., J. Transp. Res. Board*, 2002, **1802**, (1), pp. 181–189

[7] Van Lint, J.N., Zijpp, V.: 'Improving a travel-time estimation algorithm by using dual loop detectors', *Transp. Res. Rec., J. Transp. Res. Board*, 2003, **1855**, (1), pp. 41–48

[8] Sun, L., Yang, J., Mahmassani, H.: 'Travel time estimation based on piecewise truncated quadratic speed trajectory', *Transp. Res. A, Policy Pract.*, 2008, **42**, (1), pp. 173–186

[9] Ni, D., Wang, H.: 'Trajectory reconstruction for travel time estimation', *J. Intell. Transp. Syst.*, 2008, **12**, (3), pp. 113–125

[10] Mori, U., Mendiburu, A., Álvarez, M.*, et al.*: 'A review of travel time estimation and forecasting for advanced traveller information systems', *Transportmetrica A, Transp. Sci.*, 2015, **11**, (2), pp. 119–157

[11] Coifman, B.: 'Estimating travel times and vehicle trajectories on freeways using dual loop detectors', *Transp. Res. A, Policy Pract.*, 2002, **36**, (4), pp. 351–364

[12] Gazis, D.C., Herman, R., Rothery, R.W.: 'Nonlinear follow-the-leader models of traffic flow', *Oper. Res.*, 1961, **9**, (4), pp. 545–567

[13] Herman, R., Montroll, E.W., Potts, R.B.*, et al.*: 'Traffic dynamics: analysis of stability in car following', *Oper. Res.*, 1959, **7**, (1), pp. 86–106

[14] Brackstone, M., McDonald, M.: 'Car-following: a historical review', *Transp. Res. F, Traffic Psychol. Behav.*, 1999, **2**, (4), pp. 181–196

[15] Li, R., Rose, G., Sarvi, M.: 'Evaluation of speed-based travel time estimation models', *J. Transp. Eng.*, 2006, **132**, (7), pp. 540–547

[16] Yang, S., Wu, Y.J., Yin, Z.*, et al.*: 'Estimating freeway travel times using the general motors model', *Transp. Res. Rec., J. Transp. Res. Board*, 2016, **2594**, (1), pp. 83–94

[17] Revees, C.R., Rowe, J.E.: '*Genetic algorithms – principles and perspectives, a guide to GA theory*' (Kluwer Academic Publishers, Dordrecht, 2002)

[18] Rogers, S., Girolami, M.: '*A first course in machine learning*' (Chapman and Hall/CRC, 2016, 2nd edn.)

IET Intell. Transp. Syst., 2019, Vol. 13 Iss. 7, pp. 1154-1163

1163