



Traffic State Estimation Based on Eulerian and Lagrangian Observations in a Mesoscopic Modeling Framework

Aurélien Duret, Yufei Yuan

► To cite this version:

Aurélien Duret, Yufei Yuan. Traffic State Estimation Based on Eulerian and Lagrangian Observations in a Mesoscopic Modeling Framework. TRB 2017, Transportation research board annual meeting, Jan 2017, Washington DC, United States. 23p. hal-01717642

HAL Id: hal-01717642

<https://hal.science/hal-01717642>

Submitted on 16 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Traffic State Estimation Based on Eulerian and Lagrangian Observations in a Mesoscopic Modeling Framework

Aurélien Duret (corresponding author)
University of Lyon, ENTPE, IFSTTAR, LICIT, UMR-T9401
25 avenue Francois Mitterrand, 69675 Bron cedex
phone: +33 (0) 472 142 331
Aurelien.DURET@ifsttar.fr

Yufei Yuan
Delft University of Technology
Stevinweg 1, 2628CN, Delft, The Netherlands
phone: +31 (0) 15 278 6304
Y.Yuan@tudelft.nl

Paper submitted for presentation at the 96th Annual Meeting of the
Transportation Research Board, January 8-12, 2017.
July 28, 2016

5087 words + 8 figure(s) + 1 table(s) \Rightarrow 7337 ‘words’

ABSTRACT

The paper proposes a model-based framework for estimating traffic states from Eulerian (loop) and/or Lagrangian (probe) data. Lagrangian-Space formulation of the LWR model adopted as the underlying traffic model provides suitable properties for receiving both Eulerian and Lagrangian external information. Three independent methods are proposed to address Eulerian data, Lagrangian data and the combination of both, respectively. These methods are defined in a consistent framework so as to be implemented simultaneously. The proposed framework has been verified on the synthetic data derived from the same underlying traffic flow model. Strength and weakness of both data sources are discussed. Next, the proposed framework has been applied to a free-way corridor. The validity and performance have been tested using the data from a microscopic simulator.

INTRODUCTION

State of the art

Traffic state estimation (TSE) is crucial in real-time dynamic traffic management and information applications. The essence of TSE is to reproduce traffic conditions based on available observation data. One class of available estimation methods does not make use of traffic flow dynamics, but relies on basic statistics and interpolation. These are referred to as data-driven methods. Another class of estimation methods relies on dynamic traffic flow models. These are referred to as model-based methods. The focus of this article is on the latter because it potentially provides better results than the former class in non-recurrent situations (work zones, accidents, social events, etc.), regarding the monitoring-forecasting capabilities.

Model-based TSE relies on two components: a model-based component and a data assimilation algorithm. The model-based component consists of two parts: (i) a dynamic traffic flow model to predict the evolution of the state variables; and (ii) a set of observation equations relating sensor observations to the system state. Thereafter, a data-assimilation technique is adopted to combine the model predictions with the sensor observations. For example, the Kalman filter (KF) (4) and its advanced relatives, such as Extended KF (9), Unscented KF (7), Ensemble KF (10) have been widely applied in the field of traffic state estimation.

The same traffic flow model can be formulated in three two-dimensional coordinates regarding space x , time t and vehicle number n . Laval and Leclercq (5) have presented three equivalent variational formulations of the first-order traffic flow models, namely $N(x, t)$ model, $X(t, n)$ model, $T(n, x)$ model respectively, under the theory of Hamilton-Jacobi partial differential equations. Under such defined coordinate systems, sensor observations from road networks can be defined into two categories: (i) Eulerian sensing data - observations (e.g., aggregated speeds, flows) from spatially-fixed sensors (such as inductive loops, video sensors, and radar sensors) over a fixed report frequency, this type is dominating the information sources in the field of transportation research for decades; and (ii) Lagrangian sensing data - information from probe samples at a fixed time interval (such as position and speed information of individual vehicles (4), and/or probe spacing and position information (8)), this class is becoming an increasingly popular source. In literature, most of TSE applications are based on the traditional space-time (Eulerian) formulation. Aggregated traffic quantities (e.g., flows, densities or speeds (4, 9, 10)) are usually considered as system states, but no individual vehicle tracking is involved. The popularity of this formulation is due to the fact that incorporating Eulerian data is straightforward and intuitive. Recent studies have shown that a first-order (LWR) traffic flow model can be formulated and solved more efficiently and accurately in vehicle number-time (Lagrangian-time) coordinates (6). And its related Lagrangian formulation of TSE enables more accurate and efficient application of data assimilation methods, due to the solution to the mode-switching problem (traffic information travels in one direction), less non-linearity of the system model, and the nature set of observation equations to deal with Lagrangian data (11). However, the computation cost depends on the discretized platoon size (set to 1 vehicle classically) and time grid (often set around 1 second), which might be time consuming. To the best of our knowledge, none of previous research has provided a complete TSE framework for assimilating both Eulerian and Lagrangian data under a vehicle number - space (Lagrangian-space:= L-S) formulation.

Objectives and contributions

This paper presents a generic data assimilation framework based on a mesoscopic-LWR model formulated in Lagrangian-space coordinates, using both Lagrangian and Eulerian observations. The term mesoscopic is in response to the two other counterparts, since the Lagrangian-time coordinates can apply in a microscopic simulation framework and the Eulerian coordinates can accommodate in a macroscopic one. In this work, the system model is the Lagrangian-space formulation of the LWR model. It individually represents vehicles but only tracks their states at cell boundaries. We will develop algorithms and observation models to incorporate data from both Eulerian and Lagrangian sensors. And we do not apply specific data assimilation techniques; instead we try to demonstrate the sequential data assimilation concepts via reasonable assumptions. The algorithms on how to estimate network traffic states under the proposed model-based framework from the two data sources will be the main contribution of this work.

Contents of the paper

This paper is organized as follows. Section 3 presents the underlying traffic dynamics, including its formulation, solutions and properties. Section 4 describes the methodology of the proposed TSE framework, including how to assimilate Lagrangian data and combine with Eulerian data. Sections 5 and 6 illustrate the model validation and an application to a freeway corridor. Discussion and conclusions are drawn in Sections 7 and 8.

LS-LWR MODEL

This section defines the underlying process model in the state estimation framework, where the model formulation, numerical solution and its properties are discussed.

Conservation law and variational theory

This section first presents a mesoscopic formulation of the LWR model as the process model in the estimation framework. The LWR model is formulated in vehicle platoon and space (n, x) coordinates. The current mesoscopic formulation combines a vehicular description with macroscopic behavioral rules. It relaxes the temporal coordinate, and this entitles a transformation of a temporal progressing approach (e.g., in Eulerian or Lagrangian-time simulation framework) to an event-progressing approach (trigger event can be the change of time headway or pace, and/or a correction procedure based on an observation from fixed loops or probe vehicles).

The formulation follows the principle of the Hamilton-Jacobi (HJ) theory, to find an expression of the LWR model in Lagrangian-space coordinates. This model is also referred to as the T -model.

The LWR model can then be described by a hyperbolic equation:

$$\partial_x h - \partial_N(1/V(h)) = 0 \quad (1)$$

Here, h denotes the time headway. The inverse speed $1/v$ (or called pace τ) can be derived from the fundamental diagram $1/V(h)$.

Previous authors have proposed to apply variational theory in Eulerian coordinates (x, t) (1) and Lagrangian coordinates (n, t) (6). Here, we transpose the demonstration in Lagrangian-space coordinates (n, x) , following the same rationale in (6). The problem can be expressed as the Hamilton-Jacobi derived from the fundamental diagram:

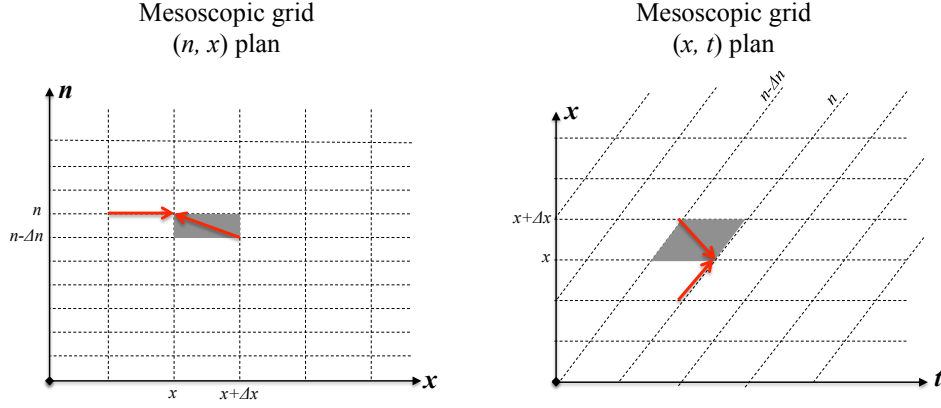


FIGURE 1 Numerical solutions in Lagrangian-space coordinates

$$\partial_x T = \frac{1}{V(\partial_N T)} \quad (2)$$

Here, the function $1/V$ represents the flux function of the problem.

Numerical solutions in Lagrangian-space coordinates

Here, a Godunov scheme is applied to solve the conservation law equation (hyperbolic equation) above with an upwind method.

This would preserve the numerical benefit of Lagrangian traffic flow models. Fig. 1 illustrates the mesoscopic numerical grid (see grey area) for the Godunov scheme. On the mesoscopic grid, the time headway is determined by:

$$h_i^{x+\Delta x} = h_i^x + \frac{\Delta x}{\Delta n} \cdot \left(\frac{1}{V(h_i^x)} - \frac{1}{V(h_{i-1}^x)} \right) \quad (3)$$

The CFL condition that guarantees the convergence of the Godunov scheme is:

$$\Delta n \geq \max_h \left| \partial_h \left(\frac{1}{V} \right) \right| \Delta x \quad (4)$$

Alternatively, the problem can also be expressed in terms of $T(n, x)$ considering the 'passage time' flux that crosses the boundary of the cell n , regarded as a variational formulation of the T -model:

$$\frac{T(n, x) - T(n, 0)}{\Delta x} = \frac{1}{V\left(\frac{T(n, x) - T(0, x)}{\Delta n}\right)} \quad (5)$$

In this expression, V depends on the fundamental diagram. Here, we consider a triangular fundamental with three parameters: the free-flow speed v_m , the maximum wave speed w and the jam density k_x . It can be expressed by:

$$\frac{1}{V(h)} = \max\left(\frac{1}{v_m}, -k_x\left(\frac{1}{wk_x} - h\right)\right) \quad (6)$$

The numerical solution to the problem is simplified as:

$$T(n, x) = \max(T(n, 0) + \frac{x}{v_m}, T(0, x + \frac{n}{k_x}) + \frac{n}{w.k_x}) \quad (7)$$

Finally, the origin $n = 0, x = 0$ could be shifted to $n - \Delta n, x - \Delta x$ and we find:

$$T(n, x) = \max(T^d, T^s), \quad (8)$$

where $T^d = T(n, x - \Delta x) + \frac{\Delta x}{v_m}$ and $T^s = T(n - \Delta n, x - \frac{\Delta n}{k_x}) + \frac{\Delta n}{w.k_x}$ represents the demand and the supply term respectively. The demand term defines the arrival time of a vehicle from upstream in non-constrained (free-flowing) conditions. The resulting passage time of a vehicle is at least equal to its arrival time but could be delayed due to the downstream conditions. Thus, the supply time provides such information in constrained (congested) conditions.

This numerical solution indicates traffic flow is divided into vehicle platoons of certain size Δn , and road stretch is discretized into spatial cells of certain length Δx . Note that, the cell length Δx in the simulation is not necessarily to be equal. The state in this formulation is the passage time $T(n, x)$ of vehicle platoons at cell boundaries. This state is always determined by the maximum of two uncorrelated terms: the demand (arrival) time and the supply time. For an elaborate description we refer to (5).

Properties

The current mesoscopic formulation is based on the notions from the variational theory. It can incorporate the numerical benefits and modeling flexibility of both Eulerian and Lagrangian-time models. Simultaneously, this formulation allows state distinction on both cell class and vehicle class, combining a vehicular description with macroscopic behavioral rules. It individually represents vehicles (platoons) but only tracks their passage times at cell boundaries. Therefore, travel times can be easily derived from the model, which is more convenient compared to other (e.g., Eulerian or Lagrangian) formulations of state estimation. This discrete model evolves state by state, with only one expression to consider all traffic conditions. Hence, it does not require memory and it is more flexible and time-efficient for data assimilation (no complex matrix inversion and multiplication). Moreover, the numerical scheme allows for long cells and cell boundaries can be located at network discontinuities only (merges, diverges, and lane-drops). In this way, the spatial discontinuities can address easily. The computation cost depends on the number of cell boundaries (x -dimension) in the network and the number of vehicles (n -dimension) to propagate during the simulation. Therefore, this would improve computational efficiency for large scale applications.

More importantly, this mesoscopic scheme is particularly convenient for data assimilation. In reality, the flow characteristics are mostly observed at fixed points (e.g., spatial fixed loop data) or along vehicle trajectories (e.g., vehicle-number fixed probe data). As discussed in literature that the Eulerian formulation is suitable for incorporating loop data and the Lagrangian-time formulation is suitable for probe data assimilation, the Lagrangian-space formulation is considered to be well-compatible for assimilating both types of observations. Because these observations are located on cell boundaries of the mesoscopic grid, which makes any traffic state estimation method convenient with this approach/formulation. This formulation can be easily coupled with any data assimilation techniques to perform state estimation. Due to the nature of the mesoscopic system model, the TSE might be not restricted to discretized mesoscopic $x - n$ grids. If we know any two boundaries in the network and an observation at a certain location or of a certain vehicle, we can generalize TSE for this specific assimilation problem.

METHODOLOGY

Methods for estimating traffic states based on loop and probe data are presented in sections 4.1 and 4.2, respectively. Next a method that combines both data sources is presented in section 4.3.

First, three definitions with respect to different traffic states are given in the following:

- an observation (o-) state is a traffic state measured by a sensor
- a background (b-) state is a state forecasted by a traffic flow model
- an analysis (a-) state is the result of an analysis procedure (or algorithm) that provides the most likely state regarding o- and b-states

TSE based on loop data

A data assimilation method using sole loop data first proposed in (3). It requires the numerical scheme to be set as follows: Δn to 1 and cells boundaries at each loop location. It considers flow and speed time series collected by loop sensors at locations $\{X_{loop}\}$ with a given frequency ΔT . Then it is implemented as a sequential procedure, for which each sequence is divided into 4 successive steps:

- Step 1: the o-state and b-state are collected and transformed
- Step 2: a Global Analysis is performed to (a-) state
- Step 3: the state of the model is updated accordingly, by adjusting arrival and supply times at cells boundaries
- Step 4: the model is run to provide a background state for the next sequence

As mentioned by the authors in (3), the update of the model is a parsimonious adjustment of the demand and/or the supply terms at cell boundaries. It has to be implemented so that the CFL stability condition is respected. The reader is referred to the paper for more details and validation.

TSE based on probe data

The data assimilation framework presented above is limited to Eulerian (loop) data while nowadays increasing amount of traffic data are collected by Lagrangian (probe) sensors. Thus a TSE estimator based on Lagrangian observations becomes essential for real applications. Probe sensors collect positions of equipped vehicles at a given time frequency. They are usually processed for providing aggregated indicators, for instance the mean speed per link. However, most of the wealth of probe data is lost during the aggregating process.

In this paper, the TSE estimator enables to assimilate positions and times without any aggregation process, which allow for using most of the details of probe data. The method is divided into 4 steps.

- Step 1: the o-state and the b-state are collected and transformed
- Step 2: Global Analysis, which consists of estimating the n -index of probe vehicles
- Step 3: the model is updated accordingly, which consists of adjusting arrival and supply times at cell boundaries of the model
- Step 4: the model is run over the next sequence to provide a new background

The two following sections elaborate steps 2 and 3, which are the keys to successfully update traffic states.

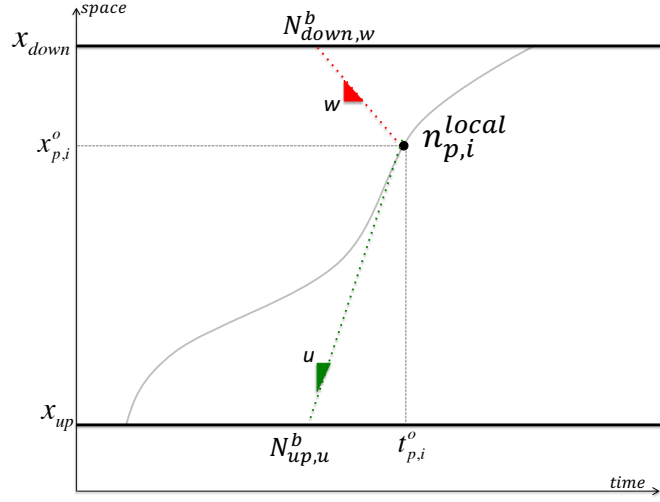


FIGURE 2 n-index estimation

1 *Focus on step 2 : Estimating the n-index of probe vehicles*

2 Let us consider the probe vehicle p that provides a set S_p of observed time-positions denoted
 3 $\{t_{p,i}^o, x_{p,i}^o\}, i \in S_p$. Simultaneously the model provides a background state $T^b(n, x)$ at cell bound-
 4 aries, from which analogous function $N^b(t, x)$ can be easily defined (T is a monotonically in-
 5 creasing function) upstream ($x = x_{up}$) and downstream ($x = x_{down}$) probe positions. $N^b(t, x_{up})$
 6 and $N^b(t, x_{down})$ can then be considered for estimating n -index of the probe based on variational
 7 principles, as illustrated in Figure 2.

$$n_{p,i}^a = \min\left(N_{up,u}^b, N_{down,w}^b + k_x \cdot (x_{down} - x_{p,i}^o)\right) \quad (9)$$

8 where

$$9 \quad \begin{cases} N_{up,u}^b = N^b\left(t_{p,i}^o - \frac{x_{up} - x_{p,i}^o}{u}, x_{up}\right) \\ N_{down,w}^b = N^b\left(t_{p,i}^o - \frac{x_{down} - x_{p,i}^o}{w}, x_{down}\right) \end{cases}$$

Equation 9 provides the n -index estimated locally (for a single time-position). At this stage, local n -index estimation could be flawed by four sources of errors: errors on the model parameters, errors on the boundary conditions, non-FIFO traffic conditions or occurrence of a traffic incident. Local errors on the estimated n -index may induce global inconsistencies on the resulting arrival/supply times. To tackle this problem, a global optimization is developed and it consists of two steps. The first step aims at building the *variational proximity matrix*, which returns the variational cost (in veh.) between each of the time-space observations from probes (with respect to the variational principles (I)). Based on that, the second step calculates the optimal n -index, denoted n^* . The optimal solution minimizes the entropy of the system while keeping a constant n -index along probe trajectories. The optimization procedure searches in the range of all possible n -indices, and this search range is defined by the minimum and maximum values from

the variational principles and the range of local n -index estimation. The entropy is defined as:

$$E(n_p^*) = \sum_i \frac{n_{p,i}^*}{n_{p,i}^{local}} \cdot \ln\left(\frac{n_{p,i}^*}{n_{p,i}^{local}}\right) \quad (10)$$

The final solution consists of the triplets $\{n_p^*, x_p^o, t_p^o\}$, where n_p^* is the optimal n -index, and t_p^o and x_p^o are the observed time and position of the probe p .

Focus on step 3: Update of arrival and supply times at cell boundaries

Once a -states are known, probe trajectories are considered as internal cell boundary conditions that are transformed into demand or supply conditions at neighboring cell boundaries. Here, we present the update of the arrival and the supply times at a cell boundary over a period P , considering that a set of probe vehicles has been analyzed.

Downstream: update of arrival times The downstream cell boundary is influenced by probe vehicles located in a time window with a length P and that moves with a free-flowing wave speed u , see Figure 3(a). Within the influencing area, each probe vehicle provides information on its upcoming arrival times. When probe vehicles travel through a cell, successive time-positions provide feasible arrival times at the downstream cell boundary. For each probe vehicle, only the latest triplet $\{n_{p,i}^o, x_{p,i}^o, t_{p,i}^o\}$ is considered for updating the arrival time at the cell downstream, as illustrated in Figure 3(b).

$$t_{a,n_p^a}^o = t_{p,i}^o + \frac{x_{down} - x_{p,i}^o}{u} \quad (11)$$

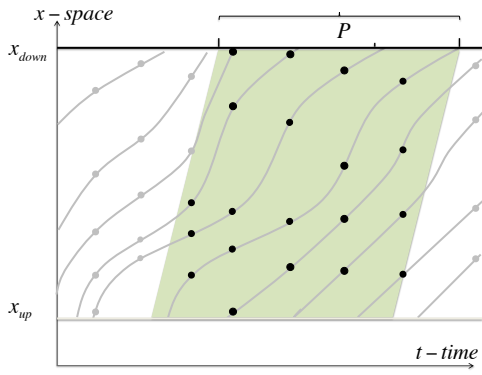
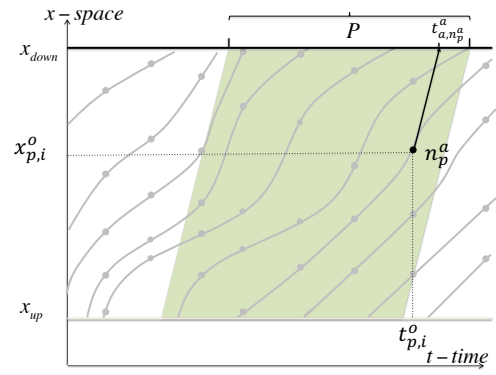
Upstream: update of supply times The upstream cell boundary is influenced by probe vehicles located in a time window with a length P and that moves with a maximum jam speed w , see Figure 3(c). For each probe vehicle, triplets $\{n_{p,i}^a, x_{p,i}^o, t_{p,i}^o\}$ are considered as internal boundary conditions to revise supply times at the cell boundary upstream. Within the influencing area, the updated supply times respect as illustrated in Figure 3(d):

$$t_{s,n_p^a+(x_{p,i}^o-x_{up}).k_x}^a = t_{p,i}^o + \frac{x_{p,i}^o - x_{up}}{w}, \forall i \in S_p \quad (12)$$

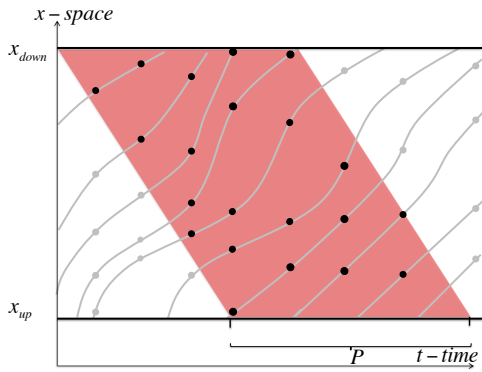
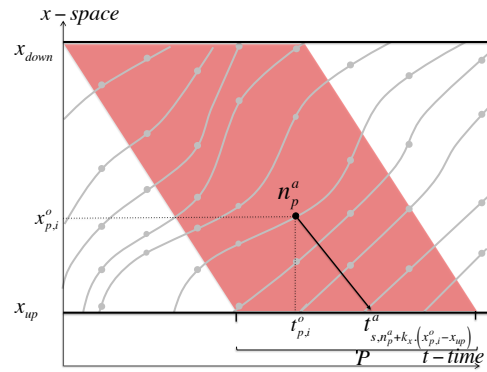
CFL condition The data assimilation process is sequential with time steps based on data time frequency ΔT . The CFL stability condition has to be respected during the sequential update of the traffic model. It requires that each cell boundary has to be updated over a time period ΔT^U , which is bounded as a wave cannot travel through a whole cell during this time period. Consequently, if $\Delta T \geq \Delta T^U$ then the updating process must proceed step by step (as described in the previous section) with a maximum time step ΔT^U .

Assimilating both loop and probe data

Loop and probe data provide information of different nature, it is therefore impossible to fuse the two data sources to perform a one-shot assimilation process. Reviewing their respective actions, the two TSE estimators act in a complementary manner. On one hand, TSE based on loop data allows for an adjustment of the flow by adding - deleting - advancing - delaying vehicles at loop

(a) Oblique window (u)

(b) Update of the arrival time

(c) Oblique window (w)

(d) Update of the supply time

sensors locations. From a physical point of view, it acts as a 'flow regulator' at cell boundaries. On the other hand, the TSE based on probe data adjusts arrival and supply times at cell boundaries considering probe trajectories as internal cell-boundary conditions. From a physical point of view, it acts as a 'travel time regulator' along cells travelled by probe vehicles.

To make the best potential use of both data, we propose first to estimate traffic states from loop data at loop sensors locations and then to estimate traffic states from probe data everywhere else. The main reason for this sequence is the following : TSE based on loop data improves the flow estimation at cell boundaries and therefore enhance the TSE based on probe data along cells. It results in a 7 steps methodology:

- Step 1: collection and transformation of the loop data and the model background states
- Step 2: Global Analysis, which consists of estimating headway-regime pairs (a-states) at each loop location
- Step 3: the model is updated accordingly (see section 4.1). At this stage, the updated model provides the best possible estimated traffic states at cell boundaries. This version of the model is considered as a new model background to be combined with probe data
- Step 4: collection and transformation of the probe data and the (updated) model background
- Step 5: Global Analysis, which consists of estimation the n-index of probes along cells
- Step 6: update of the model accordingly, by revising arrival and supply times at every cell boundaries, except those already updated during the step 3.
- Step 7: run the model over the next sequence

Here again, this sequence has to be implemented while respecting CFL stability condition mentioned in section 4.2.2.

MODEL VALIDATION

This section aims to analyze and validate the performance of the TSE methodology with loop sensors and probe sensors (separately and jointly).

Experimental validation setup

The ground truth is emulated based on a microscopic LWR model (Newell's car-following model, equivalent to the LWR model at a macroscopic scale). The model has been run on a homogeneous road stretch ($L = 2000m$, single lane) with a demand-supply scenario so that a congestion propagates through the network, see vehicle trajectories in Figure 3. A loop sensor located in the middle of the network ($x = 1000m$) collects flows and speeds with an aggregation period of 1-minute. Moreover, 5% of the vehicles are considered as probe sensors for which time-position information is reported at every 30s.

The traffic flow model is a LS-LWR model. The network is composed of 2 cells of 1000m in length, upstream and downstream of the loop sensor location. The demand-supply scenario has also been predefined with an approximative demand and a high supply so that traffic conditions are always free-flowing on the network.

TSE based on loop observation model

Figure 4 provides the estimated traffic states considering data from the loop sensor. In this figure, traffic states have been rearranged to provide travel times over the two cells. The red line provides the reference (ground truth) travel times and the blue line returns the reconstructed travel times.

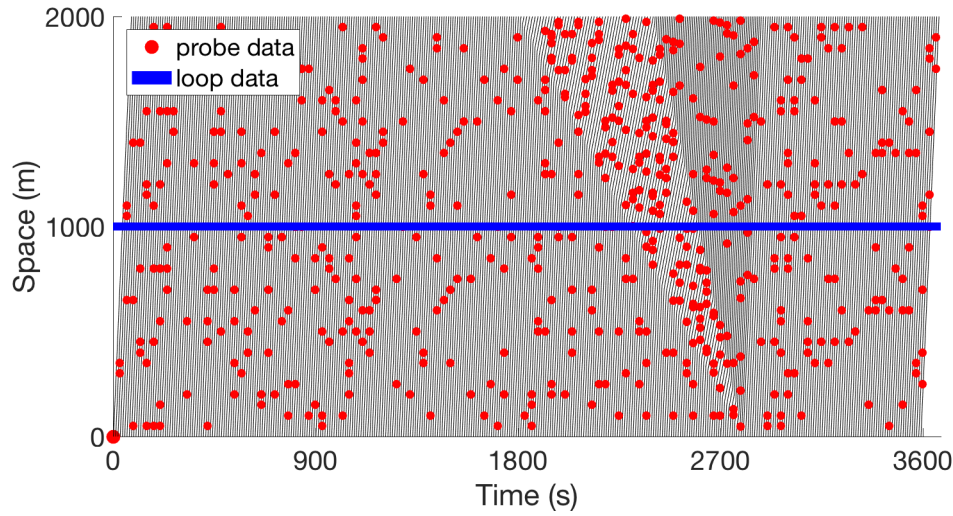


FIGURE 3 Observational model

Upstream cell Until the time $t = 2000s$, the traffic conditions are free-flowing. Between the time period $t = 2200s$ and $t = 2800s$, a congestion propagates through the upstream cell. The estimated traffic states comply with the observed travel times, which validates the ability of the TSE estimator to adjust the network supply at the loop sensor location.

Downstream cell Downstream the loop sensor, the estimated traffic states are free-flowing until the end of the simulation, whereas the observed travel times indicate that a congestion occurs. Indeed, the loop sensor data only indicate a reduced congested flow at the cell boundary, however the traffic model is unable to propagate such information toward the downstream direction but only upstream direction.

In summary, when a congestion occurs, loop sensors can estimate travel times providing that congestion states have passed over the loops. The result shows that travel times might be underestimated over the network level. And this underestimation will become significant when traffic congestion is triggered far downstream the loop sensor. We conclude that for operational purposes loop sensors have to be located as close as the triggering location of a jam/bottleneck to provide accurate estimation. In addition, the complementary information from downstream loop sensors can improve the performance of data assimilation.

TSE based on probe observation model

Figure 5 provides the estimated traffic states considering probe data only. The performance of TSE based on solo probe data provides similar performance over the two cells. It is noteworthy that TSE is very responsive as the congestion phenomenon occurs, mainly due to the probe data with a homogeneous coverage of the network both in time and space. It should also be noted that travel times are underestimated in this validation scenario due to the experimental setup. The traffic model considers a low demand versus high supply scenario. Information from probes allows for an adjustment of the supply times at the intercell boundary, but it does not rectify the underestimated flow demand (from downstream) and thus underestimate travel times. Note that the result

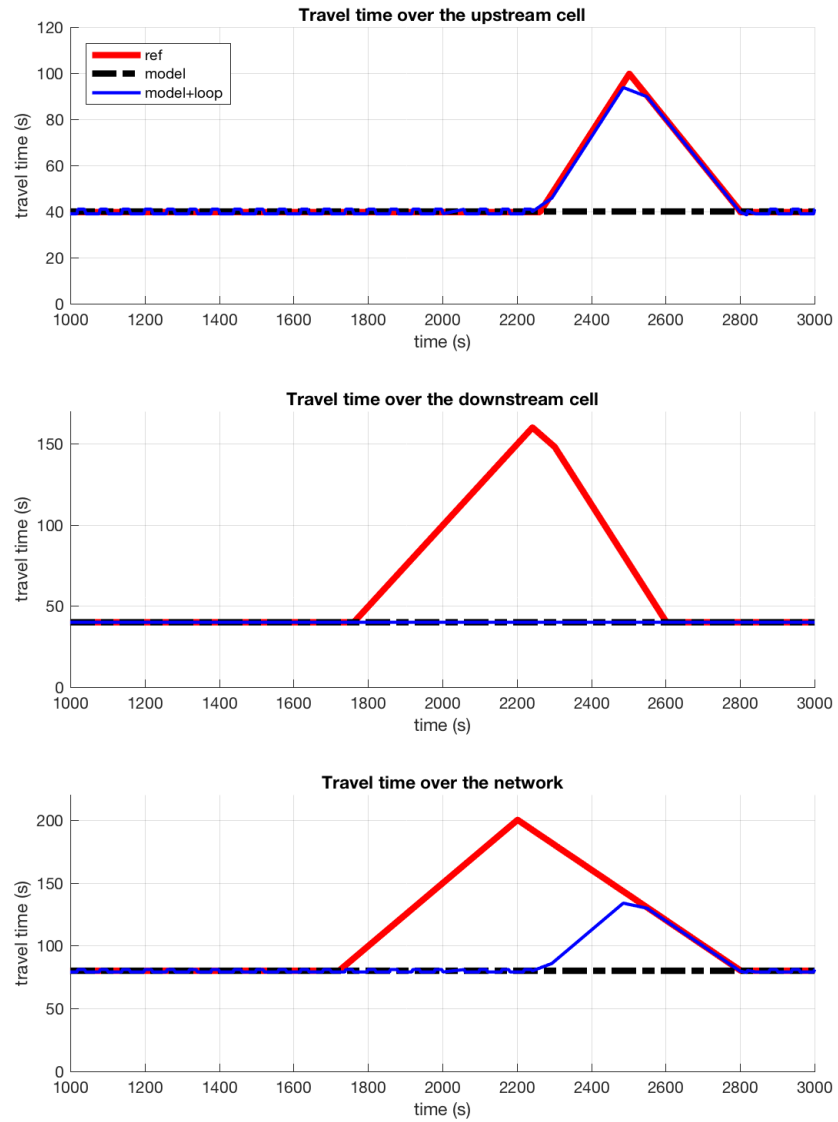


FIGURE 4 Travel time estimation from loop data assimilation

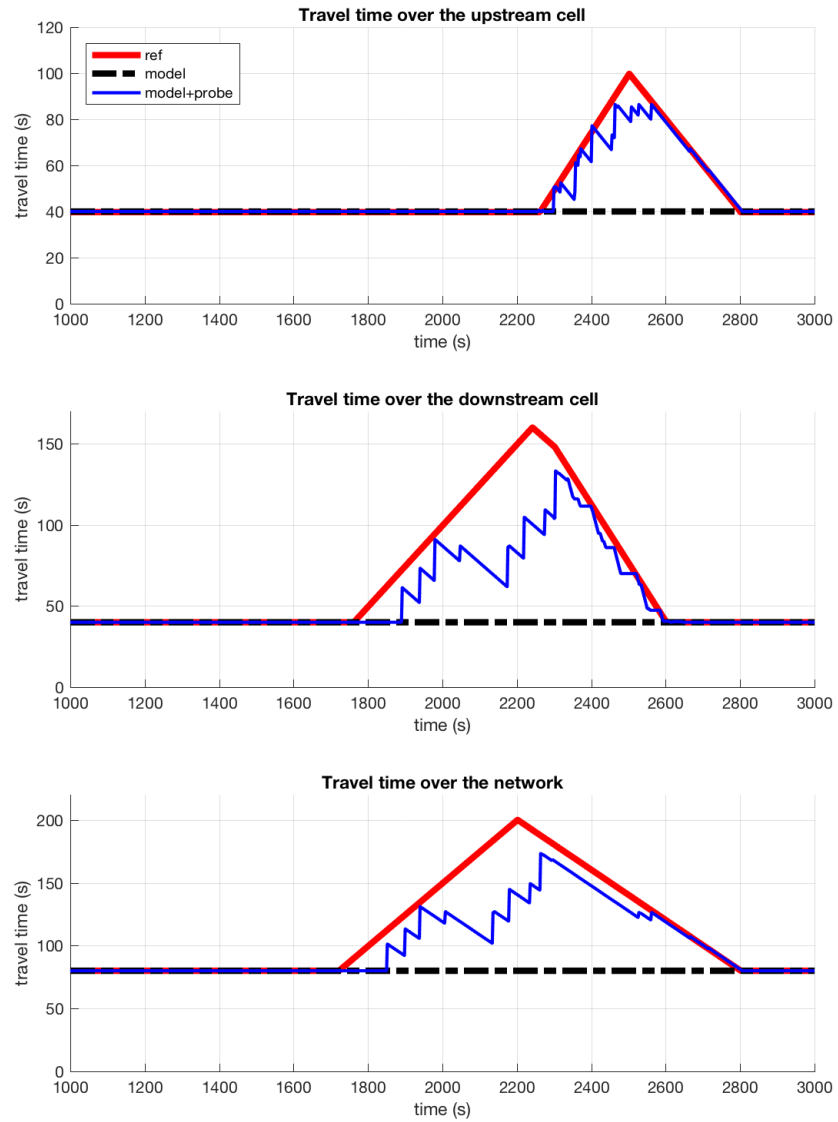


FIGURE 5 Travel time estimation from probe data assimilation

depends on the experimental setup as travel times will become overestimated if the demand from the upstream boundary is overestimated.

We conclude that for operational purposes, the knowledge of the demand at any point of the network is decisive and critical when probe data are used for estimating traffic states. The estimation can be enhanced with an accurate prior estimation of the demand; or combining probe data with loop data, as proposed in section 4.3.

TSE based on loop and probe observation model

Figure 6 provides the estimated traffic states considering loop and probe data. The results show the travel times estimated here over performs and cumulates the benefits mentioned for loop and probe observation model considered separately.

Upstream cell The performance are identical to the those provided by the loop observation model. The travel time is properly estimated and fit the ground truth travel time.

Downstream cell The performance is slightly enhanced compared to the results obtained with probe observations only. It confirms that both observations are very complementary when assimilated in the framework proposed in the paper.

APPLICATION TO A FREEWAY CORRIDOR

The previous section demonstrates the exactness of the estimator when applying to a network with FIFO conditions and homogeneous driving behavior. These assumptions are restrictive and not reflective of reality. This section aims at evaluating the performance of the estimator considering a multi-lane corridor with on- and off-ramps, with a relaxed FIFO assumption and distributed driving behavior.

Preparation of the observational model

Ground truth data have been emulated based on a microscopic traffic simulator (FOSIM (2)). This simulator is developed at the Delft University of Technology, specially designed for the detailed analysis in freeway networks. All the parameters in terms of driving behaviors have been calibrated and validated based on data from Dutch freeways. A three-lane freeway with one on-ramp and one off-ramp is designed, as illustrated in Figure 7(a) (the first 500 m as the warming-up section in Simulation, the last 1000m as the cooling-down section).

A demand-supply scenario has been built in such a way that a congestion is onset at the on-ramp. The model has been run twice: scenario 1 provides traffic conditions with only passenger cars whereas scenario 2 considers a mixed traffic condition (with 90% cars, 10% trucks). The resulting time-space diagrams and travel times are illustrated in Figure 7.

Based on FOSIM simulation results, Eulerian and Lagrangian observation models have been built. First, loop sensors have been located on the main road : loop 1 - 100m after the entrance of the network, and loop 2 - 100m upstream of on-ramp. Second, 10% equipped probe vehicles return their exact positions every 20s.

Preparation of the traffic model

The traffic model is the mesoscopic LWR model applied on a network with 7 cells : five cells for the main road (numbered from 1 to 5), one cell 6 for the off-ramp and one cell 7 for the on-ramp. Cell

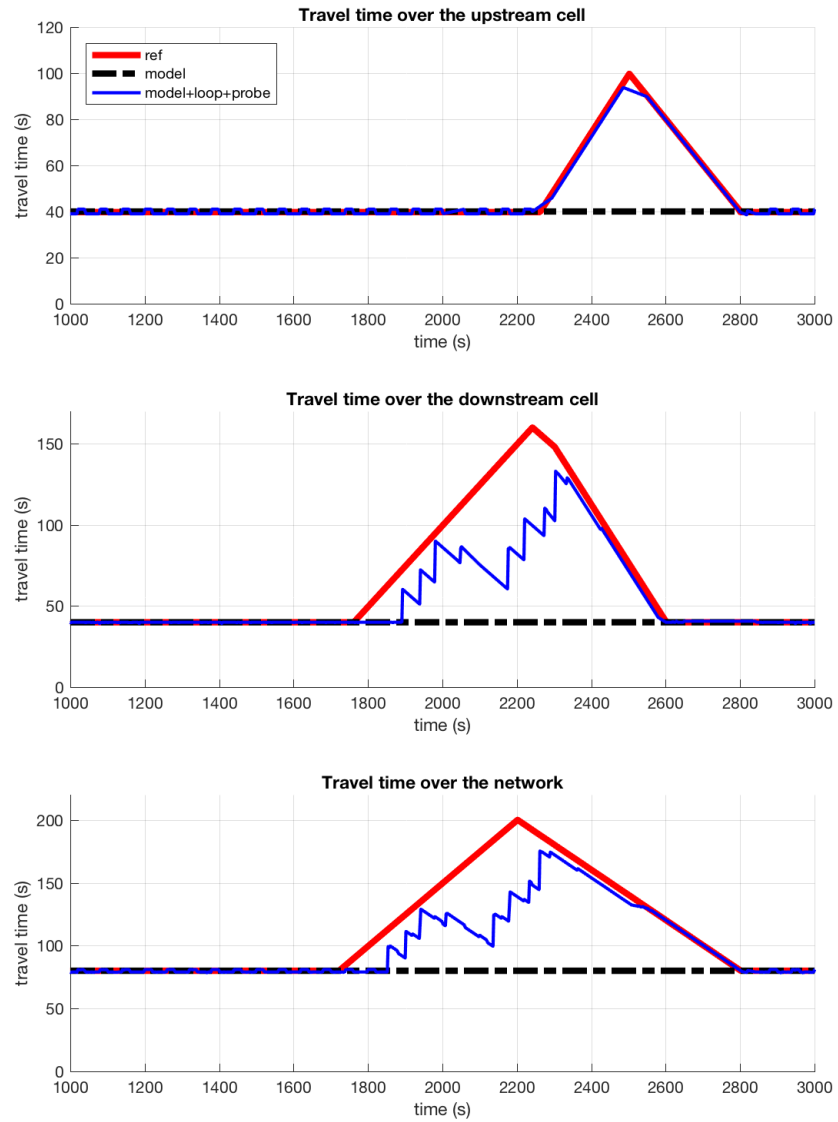
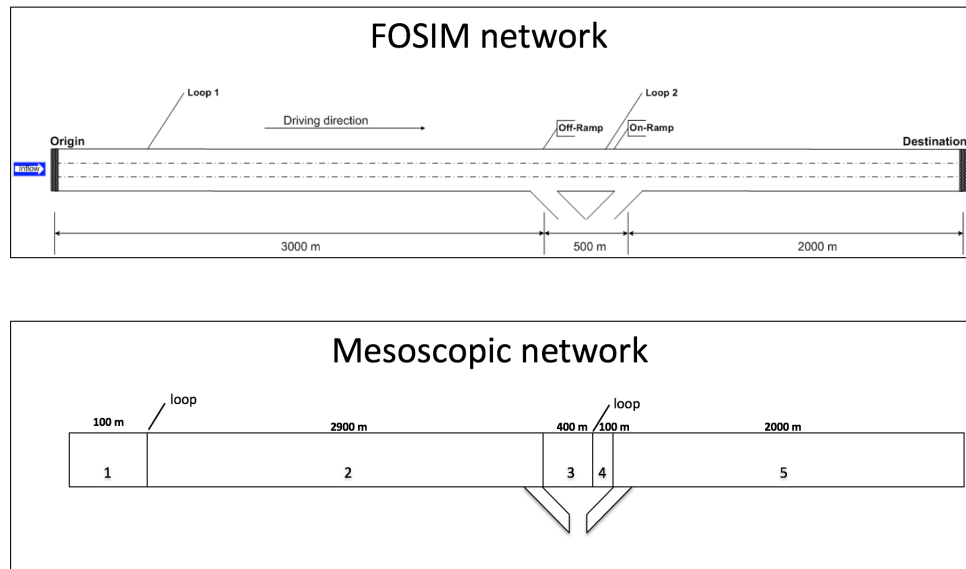
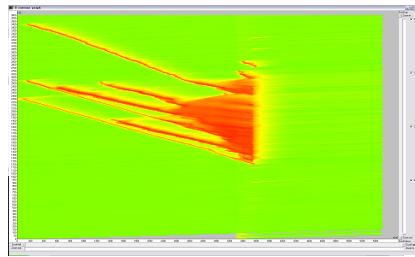


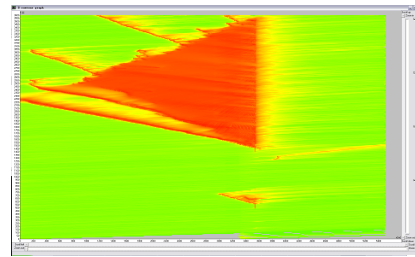
FIGURE 6 Travel time estimation from loop and probe data assimilation



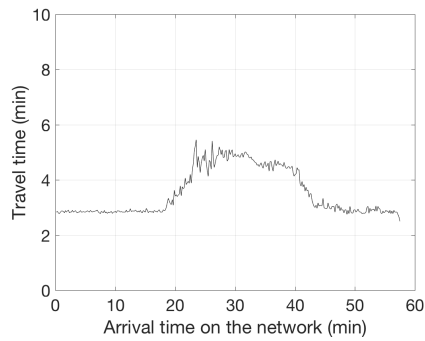
(a) Networks



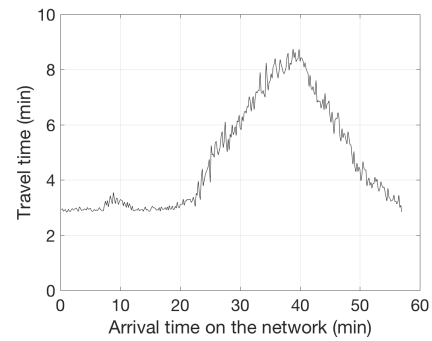
(b) time-space diagram



(c) time-space diagram



(d) travel times



(e) travel times

FIGURE 7 FOSIM observation models: cars only (a and c) and mixed traffic (b and d)

boundaries 1-2 and 3-4 are located at loop sensor locations. Boundary conditions (demand-supply) are supposed to be known approximatively and parameters of the mesoscopic LWR have been set with the following default values: $u = 110$ km/h, $w = 18$ km/h and $k_x = 150$ veh/km/lane. Results obtained from the underlying traffic model (without data assimilation) indicate that the corridor is free-flowing, with travel times stabilized around 3 mins (175s).

Results with different observation models

Three observation models have been tested : 'loop only', 'probe only', and 'combined loop and probe'. Travel time estimation based on the three observation models are illustrated in Figure 8, ground truth (in red) and default model travel times (in black) are also displayed.

During the free-flowing period, travel times are properly estimated regardless of observation models or traffic composition. However, significant differences are observed when a congestion occurs. The results analysis only focuses on the period $t = [20 - 45]min$ when the congestion is onset.

TSE based on the loop observation model underestimates travel times during the congestion period, regardless of traffic composition. This can be caused by an underestimation of the upstream demand and/or an overestimation of the supply. Loop 1 located at the entrance of the network is supposed to update the demand according to the ground truth, so the overestimation of the supply is the cause: loop 2 is located 100 meters upstream the head of the congestion, which cannot detect immediately after its onset.

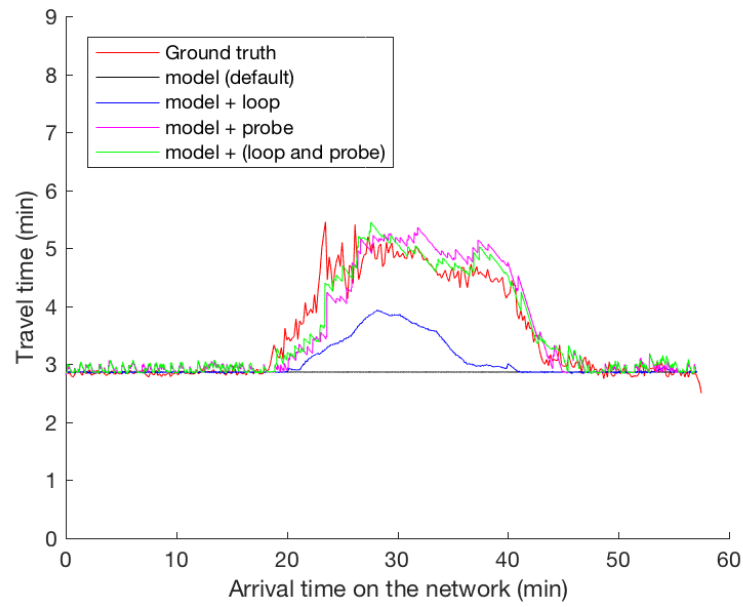
TSE based on the probe observation model presents a better performance. However, it tends to overestimate travel times. It can be caused by a poor prior estimation of demand, which skews the n -index estimation of probe vehicle and leads to poor estimation of arrival/supply times. It can also be caused by poorly calibrated traffic parameters in the traffic model and/or non FIFO observations, which is confirmed in Figure 8(b) that shows the overestimation is enhanced for a mixed traffic (ranging from $t = [30 - 40]min$). By analyzing FOSIM trajectories, it is observed that during congestion trucks are stuck on the right-most lane (over congested) while most of the cars travel faster on left-most lanes. The FIFO assumption is not fulfilled and the consequence on the performance of TSE can be explained as follows. When a probe vehicle (for instance a truck) returns its position, arrival and supply times are estimated in the (FIFO) mesoscopic model thereof. We conclude that when the characteristics of probe vehicles are distant from the mean traffic stream, this induces bias in traffic state estimation which tends to overestimate travel times.

As expected, *TSE combining loop and probe observations* outperforms the estimations considering loop and probe separately. The increase of travel times is detected immediately after the onset of congestion and the estimated travel times dynamically correspond to the ground truth.

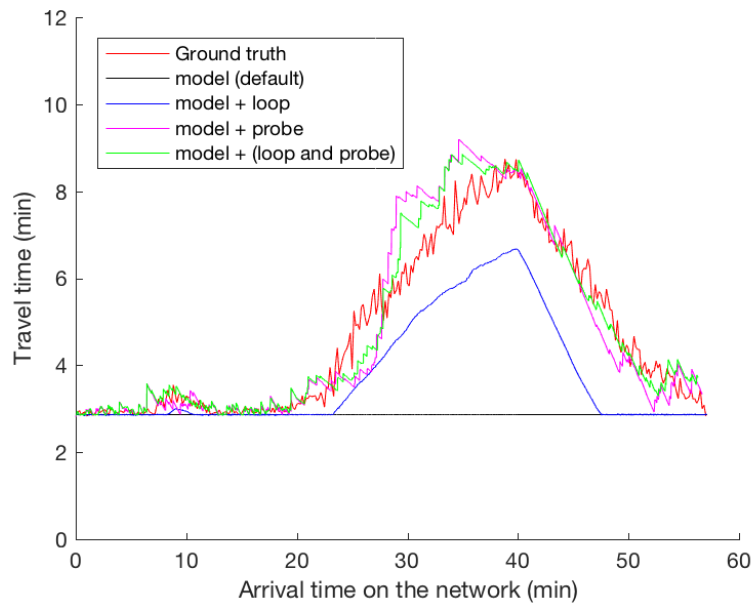
Table 1 provides three Measurements of Effectiveness (MoEs) that have been calculated over the period $t = [20 - 45]min$: mean average error (MAE) , mean average percentage error (MAPE) and mean percentage error (MPE). All the MoEs globally confirm the previous comments.

DISCUSSION

Based on the previous results, we conclude that Eulerian observations can update arrival and supply times, by adjusting the demand (flow) via adding or deleting vehicles locally. However, since loop



(a) Cars only



(b) Mixed traffic

FIGURE 8 Comparison of estimated travel times

TABLE 1 Performance of the different observational models over the indicated simulation period, Scenario 1 (cars only) And Scenario 2 (cars and trucks)

SCENARIO 1	model	model + loop	model + probe	model + (loop and probe)
RMSE (<i>s</i>)	96	69	28	21
MAPE (%)	32	24	9	6
MPE (%)	-32	-24	1.43	1
SCENARIO 2	model	model + loop	model + probe	model + (loop and probe)
RMSE (<i>s</i>)	229	92	56	42
MAPE (%)	49	22	12	10
MPE (%)	-49	-22	4	3

sensors are spatially fixed, they only catch supply information as information propagates upstream. Hence, travel time estimation might be inconsistent (see section 5.2, the case of the downstream cell). It is therefore critical to locate loop sensors at the spots of a jam and a bottleneck to provide accurate estimation.

In contrast, Lagrangian observations spread over the network in space and time. Given a reasonable resolution (10% in the validation cases), they can update both arrival times and supply times without any latency. However, two limitations exist in this method. First, the *n-index* estimation relies on the assumption of the FIFO condition, which is unrealistic. This might lead to poor estimation of travel times when probe information deviates from the average traffic conditions (see section 6.3). Second, the demand/flow at cell boundaries cannot be adjusted. A prior estimation of the demand will improve the performance of TSE with probe data (see section 5.2).

The combination of the two data sources compensates the limitations of each other. The experiment results demonstrate that TSE with data combination outperforms the estimation with a single source.

CONCLUSION

Main findings

A TSE estimator based Eulerian observations combined with a mesoscopic LWR model has been proposed and validated in (3). This paper complements the methodology with Lagrangian observations. Now both Eulerian and Lagrangian observations can be used for TSE in a unique framework.

Eulerian observations provide comprehensive observations in time and vehicle for a discrete set of locations in the network. At those locations, model states are successfully revised, which provide good performance when observations are located near the head of congestion. The update acts as a 'flow regulator' at cell boundaries by adding, deleting, advancing or delaying vehicles.

Lagrangian observations provide a homogeneous coverage of the network in time and space for a discrete set of (probe) vehicles. Probe vehicles allow for an revision of demand-supply times at neighboring cell boundary of the network. The update acts as a cell 'travel time regulator' that yields good results under the condition that the demand on the network is known. Note that the *n-index* values of probe vehicles are critical and essential in the proposed approach. The calculation of this variable is application-specific (under FIFO or non-FIFO condition) regarding estimation performance, and it is subjective for further investigation.

In the TSE framework with data combination as proposed in section 4.3, Eulerian and La-

grangian observations become highly complementary. Eulerian observations successfully update traffic states (especially the flow) at loop locations of the network while Lagrangian observations successfully update cell travel times along the network. The methodology has been verified on the synthetic data derived from the same underlying traffic flow model.

The proposed TSE framework has been applied to a freeway corridor with a relaxed FIFO condition and distributed driving behavior. The validity and performance have been tested using the ground truth from a microscopic simulator.

Further research

Future research includes, (a) to test the framework in a more general case, a realistic large-scale network with multiclass traffic, and a non-FIFO condition; (b) to test different sources of data observations, e.g., bluetooth data, information from connected vehicles; (c) to apply specific data assimilation techniques to account for model and observation reliability; (d) to study the optimal layout of Eulerian sensors and the minimal penetration of Lagrangian sensors for accurate estimation; (e) online estimation of traffic demand and important parameters in the traffic flow model; (f) traffic state prediction based on the proposed estimation framework.

ACKNOWLEDGMENT

The authors sincerely thank L. Leclercq (University of Lyon, ENTPE, IFSTTAR, LICIT, UMR-T9401) for the fruitful discussions and assistance on the methodology presented in the paper.

REFERENCES

- [1] Daganzo, C.F.: A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B: Methodological* **39**(2), 187–196 (2005)
- [2] Dijkstra, T.: Fosim (freeway operations simulation) (2012). URL <http://www.fosim.nl>
- [3] Duret, A., Leclercq, L., El Faouzi, N.E.: Data assimilation based on a mesoscopic-lwr modeling framework and loop detector data: methodology and application on a large-scale network. In: *Proceedings of the Transportation Research Board 95th Annual meeting*. Transportation Research Board, Washington, D.C. (2016)
- [4] Herrera, J.C., Bayen, A.M.: Incorporation of lagrangian measurements in freeway traffic state estimation. *Transportation Research Part B: Methodological* **44**(4), 460–481 (2010)
- [5] Laval, J., Leclercq, L.: The hamilton-jacobi partial differential equation and the three representations of traffic flow. *Transportation Research Part B: Methodological* **52**, 17–30 (2013)
- [6] Leclercq, L., Laval, J.A., Chevallier, E.: The lagrangian coordinates and what it means for first order traffic flow models. In: R. Allsop, M. Bell, B. Heydecker (eds.) *Proceedings of the 17th International Symposium on Transportation and Traffic Theory*, pp. 735–753. Elsevier, London, U.K. (2007)
- [7] Ngoduy, D.: Applicable filtering framework for online multiclass freeway network estimation. *Physica A: Statistical Mechanics and its Applications* **387**(2/3), 599–616 (2008)
- [8] Seo, T., Kusakabe, T.: Probe vehicle-based traffic state estimation method with spacing information and conservation law. *Transportation Research Part C: Emerging Technologies* **59**, 391–403 (2015)
- [9] Wang, Y., Papageorgiou, M.: Real-time freeway traffic state estimation based on extended kalman filter: A general approach. *Transportation Research Part B: Methodological* **39**(2), 141–167 (2005)
- [10] Work, D., Tossavainen, O.P., Blandin, S., Bayen, A., Iwuchukwu, T., Tracton, K.: An ensemble kalman filtering approach to highway traffic estimation using gps enabled mobile devices. In: *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 2141–2147. Cancun, Mexico (2008)
- [11] Yuan, Y., Van Lint, J.W.C., Wilson, R.E., Van Wageningen-Kessels, F.L.M., Hoogendoorn, S.P.: Real-time lagrangian traffic state estimator for freeways. *IEEE Transactions on Intelligent Transportation Systems* **13**(1), 59–70 (2012)