



HAL
open science

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. *Transportation Research Part B: Methodological*, 2017, 101, pp51-71. hal-01915206

HAL Id: hal-01915206

<https://hal.science/hal-01915206v1>

Submitted on 14 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^a, Yufei Yuan^b,

^a*University of Lyon, ENTPE, IFSTTAR, LICIT, UMR-T9401, 25 avenue François
Mitterrand, 69675 Bron cedex*

^b*Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg
1, 2628CN, Delft, The Netherlands*

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 freeway corridor. The validity has been tested using the data from a microscopic simulator, and the performance is satisfactory even for low rate of probe vehicles around 5%.

Keywords: traffic state estimation, data assimilation, LWR model, mesoscopic model, Eulerian observation, loop data, Lagrangian observation, probe data, traffic monitoring, traffic forecasting,

Email addresses: aurelien.duret@ifsttar.fr (Aurélien Duret),
y.yuan@tudelft.nl (Yufei Yuan)

1. Introduction

1.1. 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) [20, 3, 14] and its advanced relatives, such as Extended KF [29], Unscented KF [22], Ensemble KF [30] 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 Lelercq [17] 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 [14], and/or probe spacing and position information [25]), 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,

38 densities or speeds [20, 3, 14, 29, 30]) are usually considered as system states,
39 but no individual vehicle tracking is involved. The popularity of this formu-
40 lation is due to the fact that incorporating Eulerian data is straightforward
41 and intuitive. Recent studies have shown that a first-order (LWR) traffic flow
42 model [19, 24] can be formulated and solved more efficiently and accurately
43 in vehicle number-time (Lagrangian-time) coordinates [18]. And its related
44 Lagrangian formulation of TSE enables more accurate and efficient applica-
45 tion of data assimilation methods, due to the solution to the mode-switching
46 problem (traffic information travels in one direction), less non-linearity of
47 the system model, and the nature set of observation equations to deal with
48 Lagrangian data [31, 32]. However, the computation cost depends on the
49 discretized platoon size (set to 1 vehicle classically) and time grid (often set
50 around 1 second), which might be time consuming. The vehicle number -
51 space formulation inherits both the numerical benefits and modeling flexibi-
52 lity of the other two formulations.

53 Many previous studies have investigated the data assimilation problems
54 on freeway networks with both Eulerian and Lagrangian observations under
55 various modeling paradigms. Early studies in [3, 20] have applied the conserva-
56 tion law equations of traffic flow in Eulerian coordinates (LWR type [19, 24]
57 and Payne’s type [23], respectively) as the underlying traffic systems, to as-
58 similate both simulated fixed-detector data and probe-vehicle data. Falling
59 in this modeling category, the applications in [30, 14] have also performed
60 data assimilation with isolated Eulerian and Lagrangian data. Work et al.
61 [30] use a speed-based conservation law equation, while Herrera and Bayen
62 [14] use an extended LWR model in the Eulerian formulation. Alternatively,
63 the variational (HJ-PDE) formulation can be used as an equivalence to des-
64 cribe the same underlying traffic systems [2, 17]. This modeling approach is
65 considered to be much simpler to compute and numerically more accurate
66 under same conditions, compared with the conservation law approach. Ho-
67 wever, only a few studies have applied such formulations for state estimation
68 purposes. And the existing applications mainly adopt the aforementioned
69 $N(x, t)$ model, where it considers the evolution of cumulated vehicle counts.
70 For example, Claudel and Bayen [5] have applied the Hamilton-Jacobi for-
71 mulation and generalized Lax-Hopf formula [2] for data assimilation and re-
72 conciliation utilizing loop and probe data ; Newell’s three-detector model [21]
73 has also been extended for state estimation using heterogeneous data sources
74 [4, 9] : loop detector data, Bluetooth travel time data and probe GPS data,
75 respectively.

76 To the best of our knowledge, none of previous research has applied the
77 HJ-PDE formulation using the $X(t, n)$ model or the $T(n, x)$ model for traf-
78 fic data assimilation. As discussed, in comparison to the $X(t, n)$ model, the
79 simulation efficiency of the $T(n, x)$ model is independent on the discreti-
80 zed platoon size (set to 1 vehicle classically) and the spatio-temporal grid.
81 Meanwhile, observation data are located along the vehicle number - space
82 grids. Hence it is more advanced and convenient for data assimilation. This
83 article presents a complete TSE framework for assimilating both Eulerian
84 and Lagrangian data under a vehicle number - space (Lagrangian-space :=
85 L-S) formulation (namely, the $T(n, x)$ model). It inherits the classic modeling
86 approach that has been widely applied for the data assimilation problem in
87 the field of traffic management as well as meteorology, oceanography, image
88 processing, etc. [15, 16, 27]. This modeling paradigm consists of system equa-
89 tions that capture the evolution of the state vector over time, and observation
90 equations that capture the mapping of the state vector on the observations.

91 *1.2. Objectives and contributions*

92 This paper presents a generic data assimilation framework based on a
93 mesoscopic-LWR model formulated in Lagrangian-space coordinates, using
94 both Lagrangian and Eulerian observations. The term mesoscopic is in res-
95 ponse to the two other counterparts, since the Lagrangian-time coordinates
96 can apply in a microscopic simulation framework and the Eulerian coordi-
97 nates can accommodate in a macroscopic one. In this work, the system model
98 is the Lagrangian-space formulation of the LWR model. It individually repre-
99 sents vehicles but only tracks their states at cell boundaries. We will develop
100 algorithms and observation models to incorporate data from both Eulerian
101 and Lagrangian sensors, as well as considering the observation noise in both
102 data sources. And we do not apply specific data assimilation techniques (e.g.,
103 KF-based approaches) ; instead we try to demonstrate the sequential data as-
104 similation concepts via reasonable assumptions. The algorithms on how to
105 estimate network traffic states under the proposed model-based framework
106 from the two data sources will be the main contribution of this work.

107 *1.3. Contents of the paper*

108 This paper is organized as follows. Section 2 presents the underlying traf-
109 fic dynamics, including its formulation, solutions and properties. Section 3
110 describes the methodology of the proposed TSE framework, including how to
111 assimilate Eulerian data, Lagrangian data, and the combination of the two

112 sources, respectively. Sections 4 and 5 illustrate the model validation and an
 113 application to a freeway corridor. Discussion and conclusions are drawn in
 114 Sections 6 and 7.

115 2. LS-LWR model

116 This section defines the underlying process model in the state estimation
 117 framework, where the model formulation, numerical solution and its proper-
 118 ties are discussed.

119 2.1. Conservation law and variational theory

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

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

132 The LWR model can be described by a hyperbolic equation under the
 133 conservation law :

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

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

136 Previous authors have proposed to apply variational theory in Eulerian
 137 coordinates (x, t) [8] and Lagrangian coordinates (n, t) [18]. Here, we trans-
 138 pose the demonstration in Lagrangian-space coordinates (n, x) , following the
 139 same rationale in [18]. The problem can be expressed as the Hamilton-Jacobi
 140 derived from the fundamental diagram :

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

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

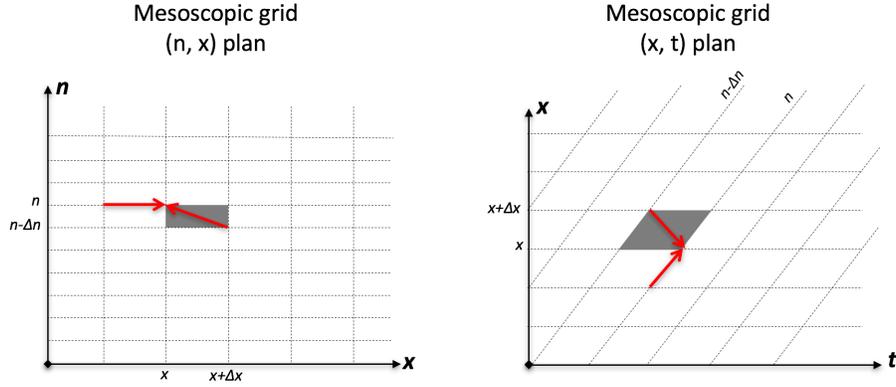


FIGURE 1: Numerical solutions in Lagrangian-space coordinates

142 *2.2. Numerical solutions in Lagrangian-space coordinates*

143 Here, a Godunov scheme [13] is applied to solve the conservation law
 144 equation (hyperbolic equation) above with an upwind method. This would
 145 preserve the numerical benefit of Lagrangian traffic flow models. Figure 1
 146 illustrates the mesoscopic numerical grid (see grey area) for the Godunov
 147 scheme. On the mesoscopic grid, the time headway is determined by :

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

148 The CFL condition that guarantees the convergence of the Godunov scheme
 149 [6] is :

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

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

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

153 In this expression, V depends on the fundamental diagram. Here, we consider
 154 a triangular fundamental with three parameters : the free-flow speed v_m ,

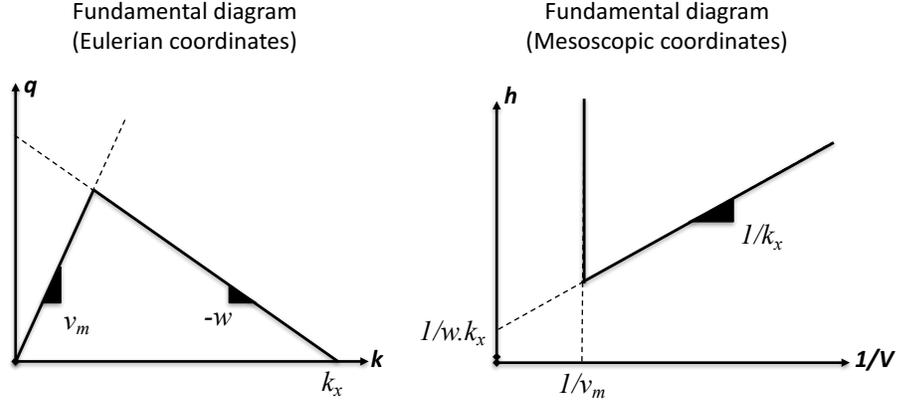


FIGURE 2: Fundamental diagrams in two coordinate systems

155 the maximum wave speed w and the jam density k_x . Figure 2 represents
 156 the resulting fundamental diagram in Eulerian coordinates (left side) and in
 157 Lagrangian-space coordinates (right side). It can be expressed by :

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

158 The numerical solution to the problem is simplified as [17] :

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

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

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

161 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
 162 the demand and the supply term, respectively. The demand term defines
 163 the arrival time of a vehicle from upstream in non-constrained (free-flowing)
 164 conditions. The resulting passage time of a vehicle is at least equal to its ar-
 165 rival time but could be delayed due to the downstream conditions. Thus, the
 166 supply time provides such information in constrained (congested) conditions.

167 This numerical solution indicates traffic flow is divided into vehicle pla-
 168 toons of certain size Δn , and road stretch is discretized into spatial cells of
 169 certain length Δx . Note that, the cell length Δx in the simulation is not
 170 necessarily to be equal. The state in this formulation is the passage time
 171 $T(n, x)$ of vehicle platoons at cell boundaries. This state is always determi-
 172 ned by the maximum of two uncorrelated terms : the demand (arrival) time
 173 and the supply time. For an elaborate description we refer to [17].

174 2.3. Properties

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

193 More importantly, this mesoscopic scheme is particularly convenient for
 194 data assimilation. In reality, the flow characteristics are mostly observed at
 195 fixed points (e.g., spatial fixed loop data) or along vehicle trajectories (e.g.,
 196 vehicle-number fixed probe data). As discussed in literature that the Eulerian
 197 formulation is suitable for incorporating loop data and the Lagrangian-time

198 formulation is suitable for probe data assimilation, the Lagrangian-space for-
199 mulation is considered to be well-compatible for assimilating both types of
200 observations. Because these observations are located on cell boundaries of
201 the mesoscopic grid, which makes any traffic state estimation method conve-
202 nient with this approach/formulation. This formulation can be easily coupled
203 with any data assimilation techniques to perform state estimation. Due to
204 the nature of the mesoscopic system model, the TSE might be not restricted
205 to discretized mesoscopic $x - n$ grids. If we know any two boundaries in the
206 network and an observation at a certain location or of a certain vehicle, we
207 can generalize TSE for this specific assimilation problem.

208 **3. Data assimilation methodology**

209 Traffic flow prediction is an initial/boundary value problem, where a traf-
210 fic model forecasts the evolution of traffic states on the network given ini-
211 tial/boundary states. Such data assimilation model-based approaches have
212 been developed and widely applied in other fields, notably for atmospheric
213 modeling and forecast [11, 27]. Because predictions become more accurate
214 when uncertainties of initial values are reduced, the data assimilation pro-
215 blem first combines observed data and the "first guess" provided by the model
216 to estimate the mostly possible initial states at observation locations, also
217 called analysis states. When the model is updated accordingly, it becomes
218 ready to forecast the evolution of the system. [15] presents the typical data
219 assimilation framework for atmospheric modeling (see chapter 5), where a
220 sequential loop is run every time new data are observed. The composition of
221 the loop returns an overarching theory of sequential data assimilation with
222 four fundamental steps :

- 223 — Step 1 : the transformation operation makes observed data and model
224 predictions comparable
- 225 — Step 2 : the global analysis provides the analysis states
- 226 — Step 3 : the model is updated accordingly
- 227 — Step 4 : the model forecasts the system evolution, by propagating
228 information from data rich to data poor areas.

229 In the paper, methods for estimating traffic states based on loop and
230 probe data are presented in sections 3.1 and 3.2, respectively. Then a method
231 that combines both data sources is presented in section 3.3.

232 Three definitions with respect to different traffic states are given in the
233 following :

- 234 — an observation (o-) state is a traffic state measured by a sensor
- 235 — a background (b-) state is a state forecasted by a traffic flow model
- 236 — an analysis (a-) state is the result of an analysis procedure (or algo-
- 237 rithm) that provides the most likely state regarding o- and b-states

238 3.1. TSE based on loop data

239 3.1.1. Reminder on the existing methodology

240 A data assimilation method using sole loop data first proposed in [12],
 241 has been validated on synthetic scenarios and tested on a large-scale network.
 242 Here the basic concepts are reviewed and the reader is referred to the paper
 243 for more details.

244 It requires the numerical scheme to be set as follows : Δn to 1 and cells
 245 boundaries at each loop location. It considers flow and speed time series
 246 collected by loop sensors at locations $\{X_{loop}\}$ with a given frequency ΔT .
 247 Then it is implemented as a sequential procedure, for which each sequence
 248 is divided into 4 successive steps (see Figure 3).

- 249 — Step 1 : the o-state and b-state are collected and transformed
- 250 — Step 2 : a Global Analysis is performed to (a-) state
- 251 — Step 3 : the state of the model is updated accordingly, by adjusting
- 252 arrival and supply times at cells boundaries
- 253 — Step4 : the model is run to provide a background state for the next
- 254 sequence

255 As mentioned by the authors in [12], the update of the model has to
 256 be implemented so that the CFL stability condition is respected [6]. The
 257 solution proposed is a parsimonious adjustment of the demand and/or the
 258 supply terms at cell boundaries.

259 3.1.2. Focus on step 3 : Update of arrival and supply times at cell boundaries

260 The a-state proposed by the Global Analysis consists of a regime r^a and a
 261 headway h^a , at observation location and over the period P . $8 (2^3)$ situations
 262 can be met (see Table 1) and the model needs to be updated accordingly.

263 The update consists of (i) adding, deleting, advancing or delaying vehicles
 264 at the cell boundaries and (ii) then updating the passing times of vehicles
 265 accordingly. In this paper, the update is identical to the one presented in [12]
 266 except that the passing times are managed slightly differently to be better
 267 suited for combining with the assimilation of probe data (see the next sec-
 268 tion) while keeping the same performance.

269

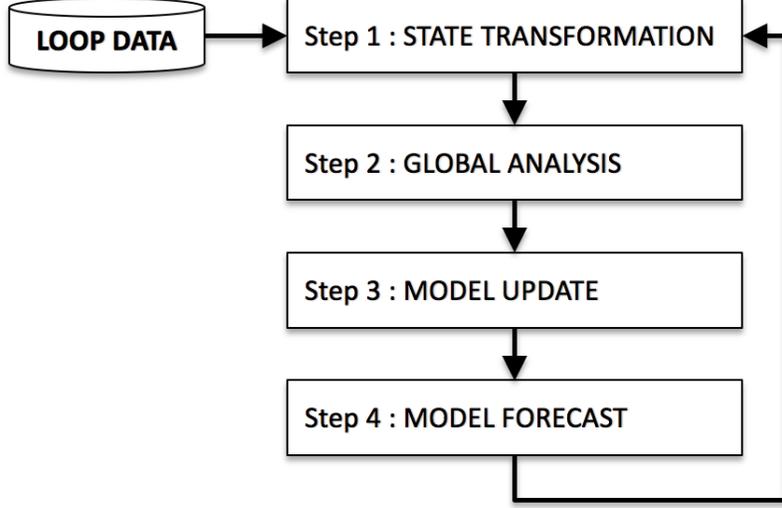


FIGURE 3: TSE based on loop data : methodology

TABLE 1: Summary of the 8 states combinations of a-state and b-state, source [12]

	$r^b = 0$	$r^b = 1$
$r^a = 0$	$h^a > h^b \dots\dots\dots(1)$	$h^a > h^b \dots\dots\dots(5)$
	$h^a \leq h^b \dots\dots\dots(2)$	$h^a \leq h^b \dots\dots\dots(6)$
$r^a = 1$	$h^a > h^b \dots\dots\dots(7)$	$h^a > h^b \dots\dots\dots(3)$
	$h^a \leq h^b \dots\dots\dots(8)$	$h^a \leq h^b \dots\dots\dots(4)$

To do so the passing times are not updated directly. The method updates the demand and the supply terms at cell boundaries over the period P

If the analysis regime is free-flowing (combination 3, 4, and 7 in Table 1) :

$$\begin{cases} T^s(n, X) = -\infty \\ T^d(n, X) = T(n-1, X) + h^a \end{cases} \quad (9)$$

270
271

If the analysis regime is congested (combination 1, 2 and 6 in Table 1) :

$$\begin{cases} T^s(n, X) = T(n-1, X) + h^a \\ T^d(n, X) = -\infty \end{cases} \quad (10)$$

272 Note that combinations 5 and 8 in Table 1 correspond to errors on both
 273 the local demand and supply. No update is required and the problem has to
 274 be addressed at global level (demand, model parameters, etc.).

275 3.2. TSE based on probe data

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

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

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

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

296 3.2.1. Focus on step 2 : Estimating the n -index of probe vehicles

297 Let us consider the probe vehicle p that provides a set S_p of observed time-
 298 positions denoted $\{t_{p,i}^o, x_{p,i}^o\}, i \in S_p$. Simultaneously the model provides a
 299 background state $T^b(n, x)$ at cell boundaries, from which analogous function
 300 $N^b(t, x)$ can be easily defined (T is a monotonically increasing function)
 301 upstream ($x = x_{up}$) and downstream ($x = x_{down}$) probe positions. $N^b(t, x_{up})$
 302 and $N^b(t, x_{down})$ can then be considered for estimating local n -index of the
 303 probe based on variational principles applied for the three-detector problem
 304 [7, 8, 21] as illustrated in Figure 5.

$$n_{p,i}^{local} = \min\left(N_{up,u}^b, N_{down,w}^b + k_x \cdot (x_{down} - x_{p,i}^o)\right) \quad (11)$$

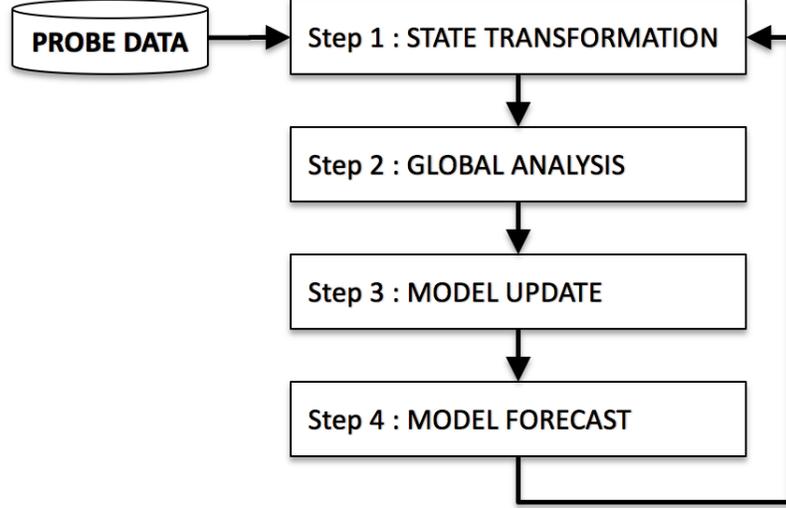


FIGURE 4: TSE based on probe data : methodology

305 where

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

Equation 11 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 [8]). Based on that, the second step calculates the optimal n -index, denoted n^* . The optimal solution minimizes the entropy of the system [1, 28] 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

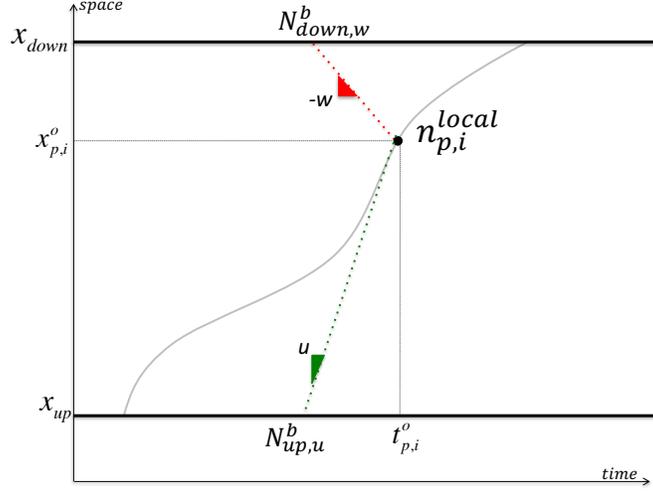


FIGURE 5: n-index estimation

entropy is defined according to [26] :

$$E(n_p^*) = \sum_i \frac{n_{p,i}^*}{n_{p,i}^{local}} \cdot \ln\left(\frac{n_{p,i}^*}{n_{p,i}^{local}}\right) \quad (12)$$

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

310 *3.2.2. Focus on step 3 : Update of arrival and supply times at cell boundaries*

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

316 *Downstream : update of arrival times.* The downstream cell boundary is in-
 317 fluenced by probe vehicles located in a time window with a length P and
 318 that moves with a free-flowing wave speed u , see Figure 6.

319 Within the influencing area, each probe vehicle provides information on its
 320 upcoming arrival times. When probe vehicles travel through a cell, successive

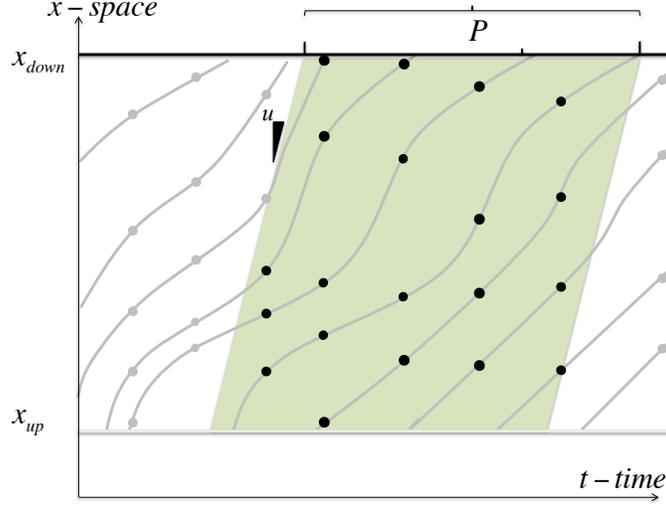


FIGURE 6: Oblique window for updating the demand term at the downstream cell boundary

321 time-positions provide feasible arrival times at the downstream cell boundary.
 322 For each probe vehicle, only the latest triplet $\{n_p^a, x_{p,i}^o, t_{p,i}^o\}$ is considered for
 323 updating the arrival time at the cell downstream, as illustrated in Figure 7.

$$t_{a,n_p^a}^a = t_{p,i}^o + \frac{x_{down} - x_{p,i}^o}{u} \quad (13)$$

324 *Upstream : update of supply times.* The upstream cell boundary is influenced
 325 by probe vehicles located in a time window with a length P and that moves
 326 with a maximum jam speed w , see Figure 8.

327 For each probe vehicle, triplets $\{n_{p,i}^a, x_{p,i}^o, t_{p,i}^o\}$ are considered as internal
 328 boundary conditions to revise supply times at the cell boundary upstream.
 329 Within the influencing area, the updated supply times respect as illustrated
 330 in Figure 9 :

$$t_{s,n_p^a+(x_{p,i}^o-x_{up})\cdot k_x}^a = t_{p,i}^o + \frac{x_{p,i}^o - x_{up}}{w}, \forall i \in S_p \quad (14)$$

331 **CFL condition** The data assimilation process is sequential with time
 332 steps based on data time frequency ΔT . The CFL stability condition has to
 333 be respected during the sequential update of the traffic model. It requires

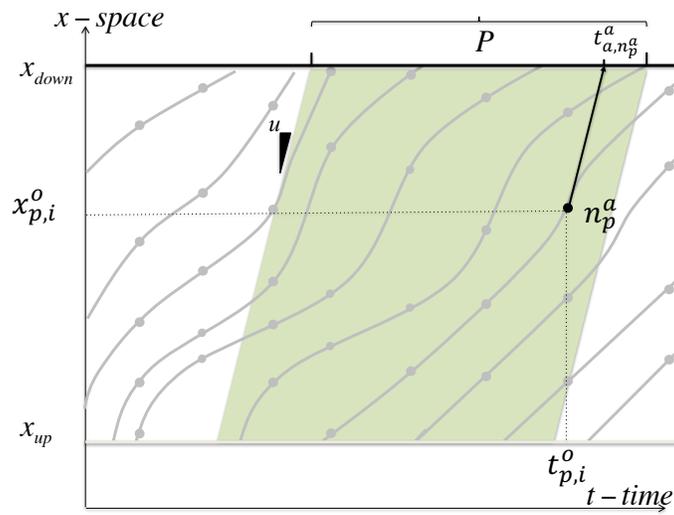


FIGURE 7: Update of the arrival time at the downstream cell boundary

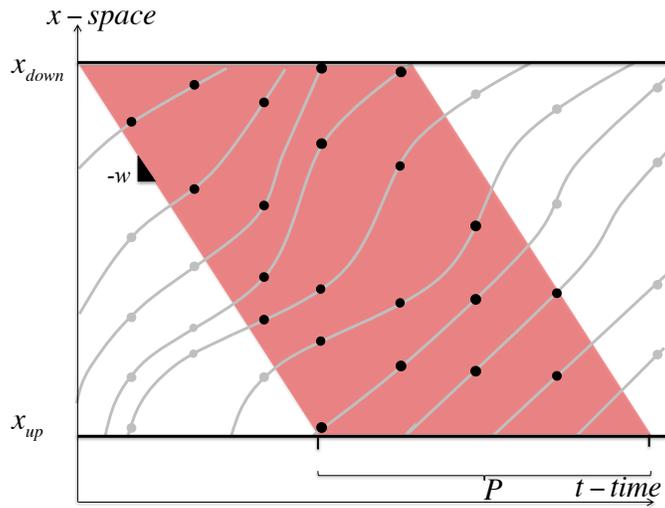


FIGURE 8: Oblique window for updating the supply term at the upstream cell boundary

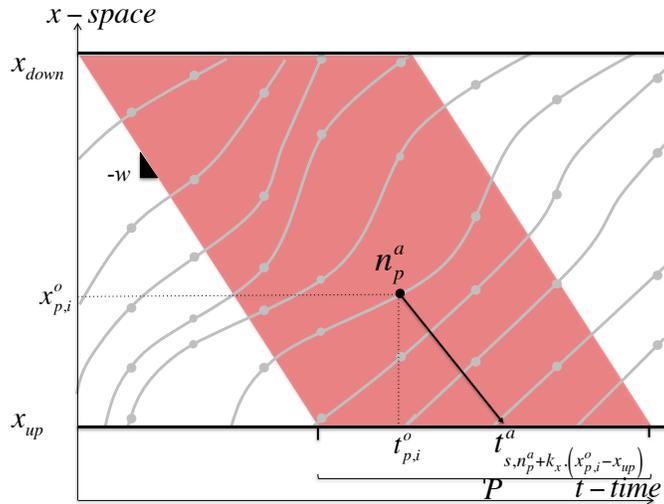


FIGURE 9: Update of the supply times at the upstream cell boundary

334 that each cell boundary has to be updated over a time period ΔT^U , which
 335 is bounded as a wave cannot travel through a whole cell during this time
 336 period. Consequently, if $\Delta T \geq \Delta T^U$ then the updating process must proceed
 337 step by step (as described in the previous section) with a maximum time
 338 step ΔT^U .

339 3.3. Assimilating both loop and probe data

340 Loop and probe data provide information of different nature, it is there-
 341 fore impossible to fuse the two data sources to perform a one-shot assimilation
 342 process. Reviewing their respective actions, the two TSE estimators act in a
 343 complementary manner. On one hand, TSE based on loop data allows for an
 344 adjustment of the flow by adding - deleting - advancing - delaying vehicles at
 345 loop sensors locations. From a physical point of view, it acts as a 'flow regu-
 346 lator' at cell boundaries. On the other hand, the TSE based on probe data
 347 adjusts arrival and supply times at cell boundaries considering probe trajec-
 348 tories as internal cell-boundary conditions. From a physical point of view, it
 349 acts as a 'travel time regulator' along cells travelled by probe vehicles.

350 To make the best potential use of both data, we propose first to estimate
 351 traffic states from loop data at loop sensors locations and then to estimate
 352 traffic states from probe data everywhere else. The main reason for this se-

353 quence is the following : TSE based on loop data improves the flow estimation
354 at cell boundaries and therefore enhance the TSE based on probe data along
355 cells. It results in a 7 steps methodology illustrated on Figure 10 :

- 356 — Step 1 : collection and transformation of the loop data and the model
357 background states
- 358 — Step 2 : Global Analysis, which consists of estimating headway-regime
359 pairs (a-states) at each loop location
- 360 — Step 3 : the model is updated accordingly (see section 3.1). At this
361 stage, the updated model provides the best possible estimated traffic
362 states at cell boundaries. This version of the model is considered as a
363 new model background to be combined with probe data
- 364 — Step 4 : collection and transformation of the probe data and the (up-
365 dated) model background
- 366 — Step 5 : Global Analysis, which consists of estimation the n-index of
367 probes along cells
- 368 — Step 6 : update of the model accordingly, by revising arrival and supply
369 times at every cell boundaries, except those already updated during
370 the step 3.
- 371 — Step 7 : run the model over the next sequence

372 Here again, this sequence has to be implemented respecting CFL stability
373 condition mentioned in section 3.2.2.

374 4. Model validation

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

377 4.1. Experimental validation setup

378 The ground truth is emulated based on a microscopic LWR model (Ne-
379 well's car-following model [21], equivalent to the LWR model at a macroscopic
380 scale). The model has been run on a homogeneous road stretch ($L = 2000m$,
381 single lane) with a demand-supply scenario so that a congestion propagates
382 through the network, see vehicle trajectories in Figure 11. A loop sensor lo-
383 cated in the middle of the network ($x = 1000m$) collects flows and speeds
384 with an aggregation period of 1-minute. Moreover, 10% of the vehicles are
385 considered as probe sensors for which time-position information is reported
386 at every 30s.

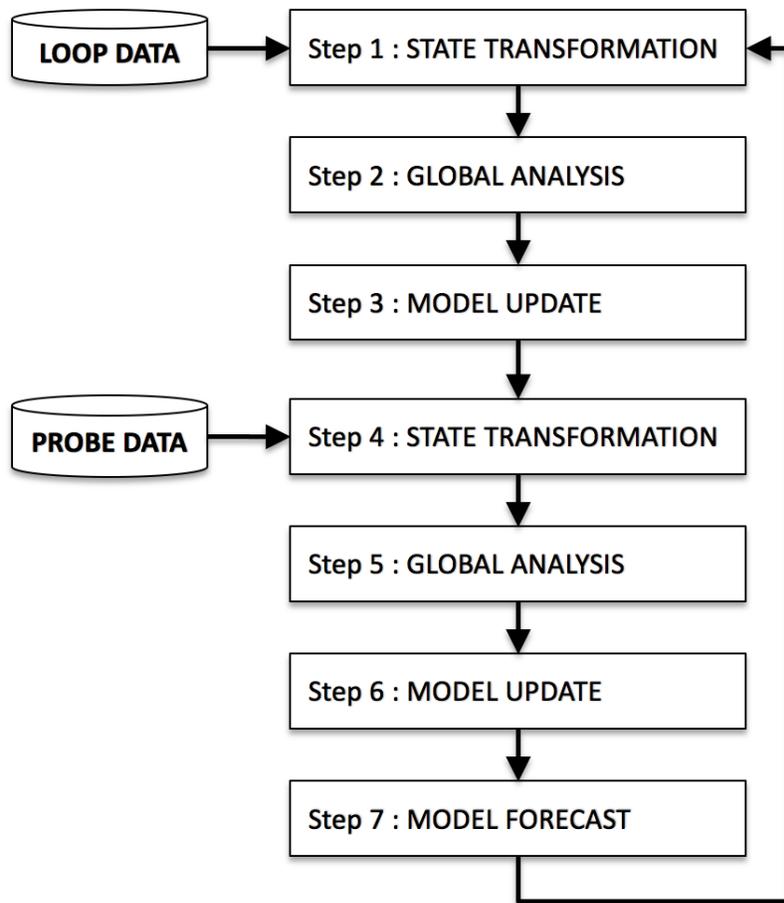


FIGURE 10: TSE based on loop and probe data : methodology

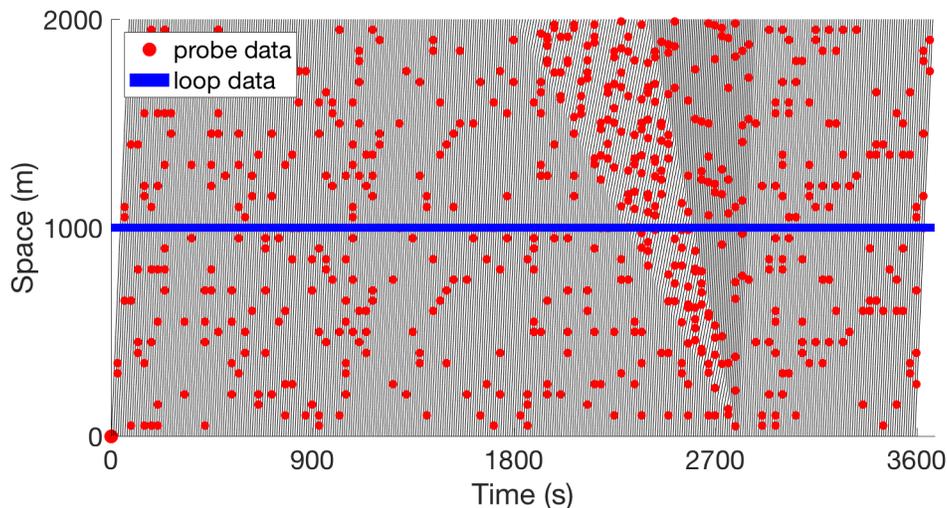


FIGURE 11: Observational model

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

392 4.2. TSE based on loop observation model

393 Figure 12 provides the estimated traffic states considering data from the
 394 loop sensor. In this figure, traffic states have been rearranged to provide
 395 travel times over the two cells. The red line provides the reference (ground
 396 truth) travel times and the blue line returns the reconstructed travel times.

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

402 *Downstream cell.* Downstream the loop sensor, the estimated traffic states
 403 are free-flowing until the end of the simulation, whereas the ground truth
 404 indicate that a congestion occurs. Indeed, the loop sensor data only indicate

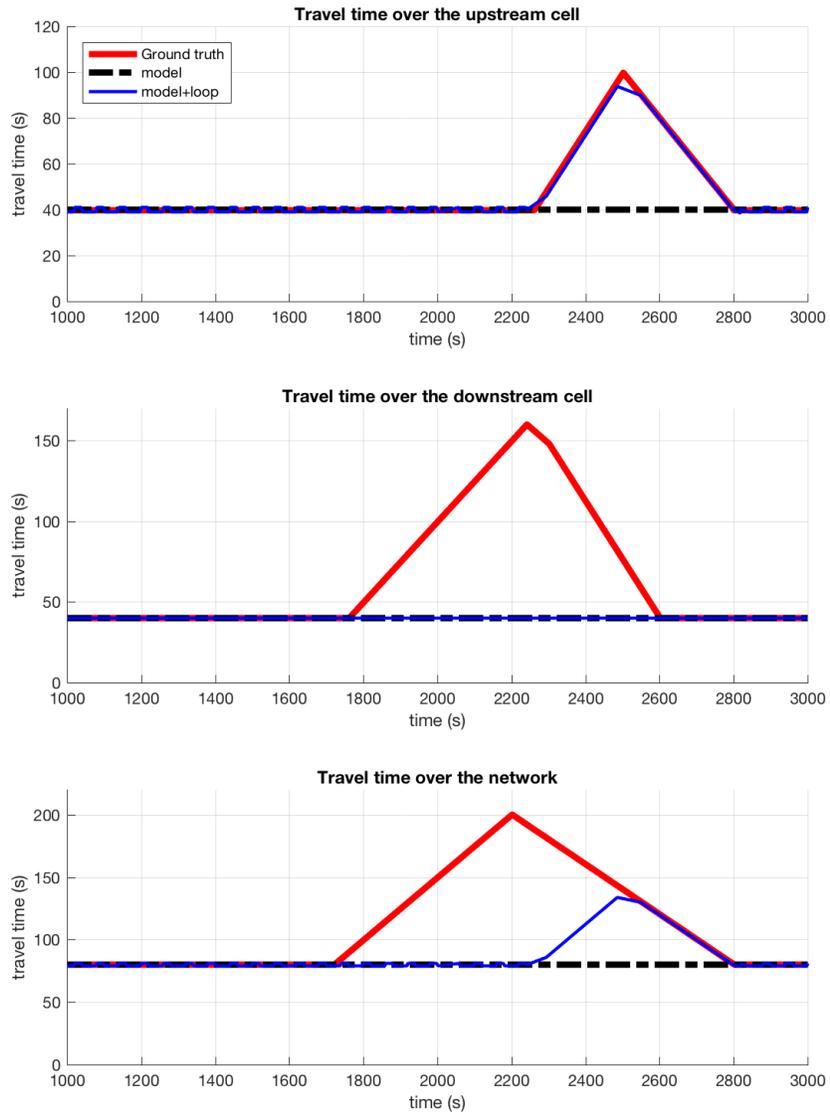


FIGURE 12: Travel time estimation from loop data assimilation

405 a reduced congested flow at the cell boundary, however the traffic model is
406 unable to propagate such information toward the downstream direction but
407 only upstream direction.

408 In summary, when a congestion occurs, loop sensors can estimate tra-
409 vel times providing that congestion states have passed over the loops. The
410 result shows that travel times might be underestimated over the network le-
411 vel. And this underestimation will become significant when traffic congestion
412 is triggered far downstream the loop sensor. We conclude that for opera-
413 tional purposes loop sensors have to be located as close as the triggering
414 location of a jam/bottleneck to provide accurate estimation. In addition, the
415 complementary information from downstream loop sensors can improve the
416 performance of data assimilation.

417 *4.3. TSE based on probe observation model*

418 Figure 13 provides the estimated traffic states considering probe data
419 only. The performance of TSE based on sole probe data provides similar per-
420 formance over the two cells. It is noteworthy that TSE is very responsive
421 as the congestion phenomenon occurs, mainly due to the probe data with
422 a homogeneous coverage of the network both in time and space. It should
423 also be noted that travel times are underestimated in this validation scenario
424 due to the experimental setup. The traffic model considers a low demand
425 versus high supply scenario. Information from probes allows for an adjust-
426 ment of the supply times at the intercell boundary, but it does not rectify
427 the underestimated flow demand (from downstream) and thus underestimate
428 travel times. Note that the result depends on the experimental setup as travel
429 times will become overestimated if the demand from the upstream boundary
430 is overestimated.

431 We conclude that for operational purposes, the knowledge of the demand
432 at any point of the network is decisive and critical when probe data are
433 used for estimating traffic states. The estimation can be enhanced with an
434 accurate prior estimation of the demand ; or combining probe data with loop
435 data, as proposed in section 3.3.

436 *4.4. TSE based on loop and probe observation model*

437 Figure 14 provides the estimated traffic states considering both loop and
438 probe data. The results show the travel time estimation here outperforms
439 and cumulates the benefits mentioned for loop and probe observation model
440 considered separately.

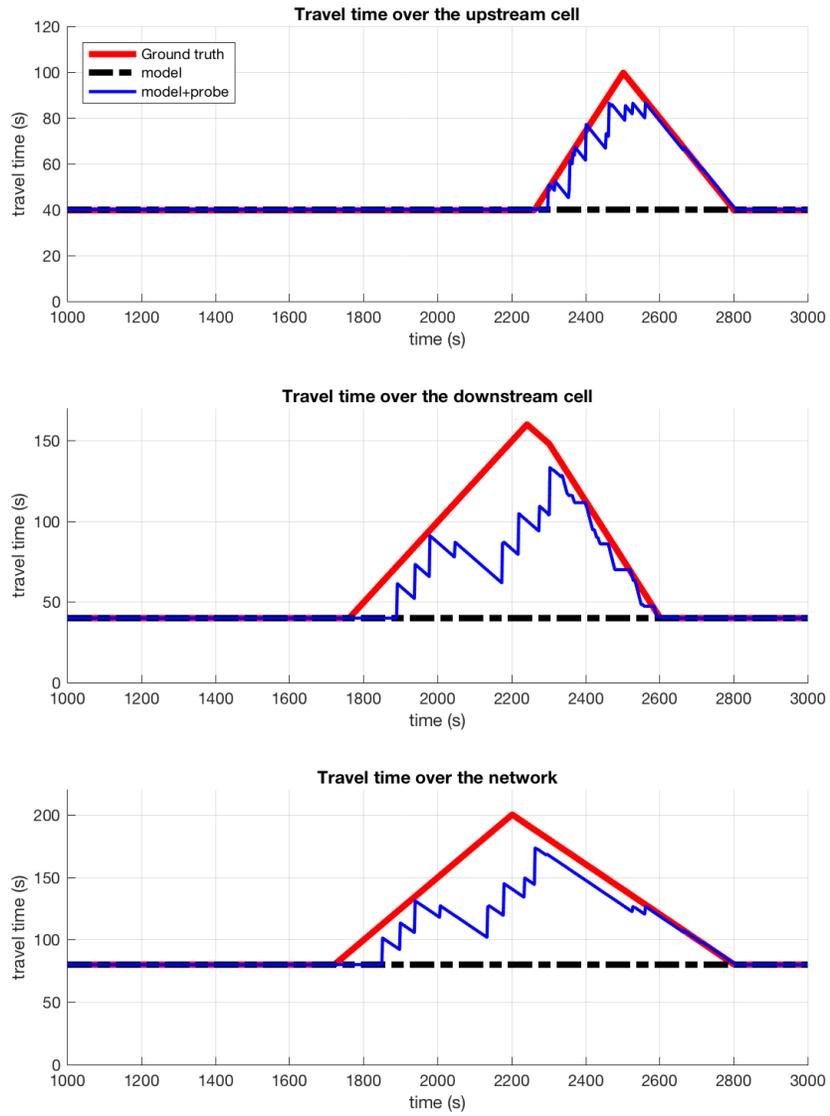


FIGURE 13: Travel time estimation from probe data assimilation

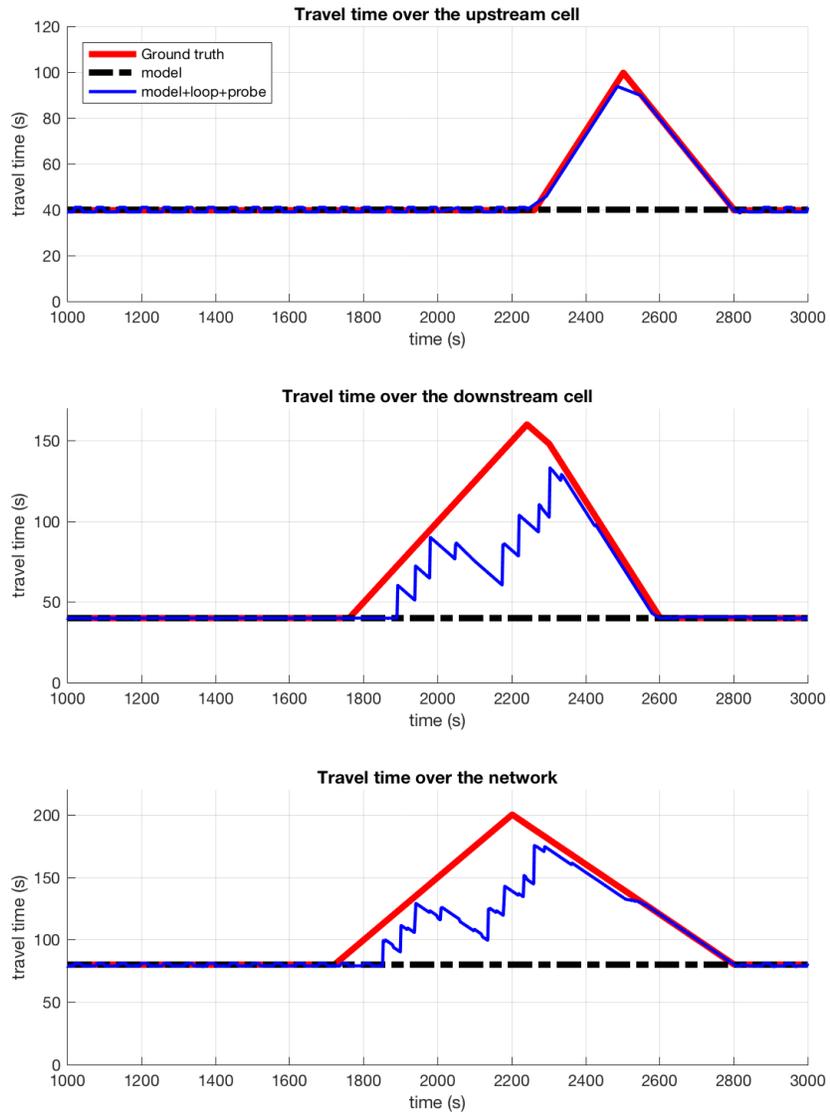


FIGURE 14: Travel time estimation from loop and probe data assimilation

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

444 *Downstream cell.* The performance is slightly enhanced compared to the re-
 445 sults obtained with probe observations only. It confirms that both observa-
 446 tions are very complementary when assimilated in the framework proposed
 447 in the paper.

448 Table 2 provides three Measurements of Effectiveness (MoEs) that have
 449 been calculated over the congested period = [30 - 45]min : root mean square
 450 error (RMSE) , mean absolute percentage error (MAPE) and mean percent-
 451 age error (MPE). All the MoEs globally confirm the previous comments.

TABLE 2: Scenario with a homogeneous stretch of road

	model	model + loop	model + probe	model + (loop and probe)
RMSE (s)	229	58	28	28
MAPE (%)	34	25	12	11
MPE (%)	-34	-25	-12	-11

452 5. Application to a freeway corridor

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

459 5.1. Preparation of the observational model

460 Ground truth data have been emulated based on a microscopic traffic
 461 simulator (FOSIM [10]). This simulator is developed at the Delft University of
 462 Technology, specially designed for the detailed analysis in freeway networks.
 463 All the parameters in terms of driving behaviors have been calibrated and
 464 validated based on data from Dutch freeways. A three-lane freeway with
 465 one on-ramp and one off-ramp is designed, as illustrated in Figure 15 (the
 466 first 500 m as the warming-up section in Simulation, the last 1000m as the
 467 cooling-down section).

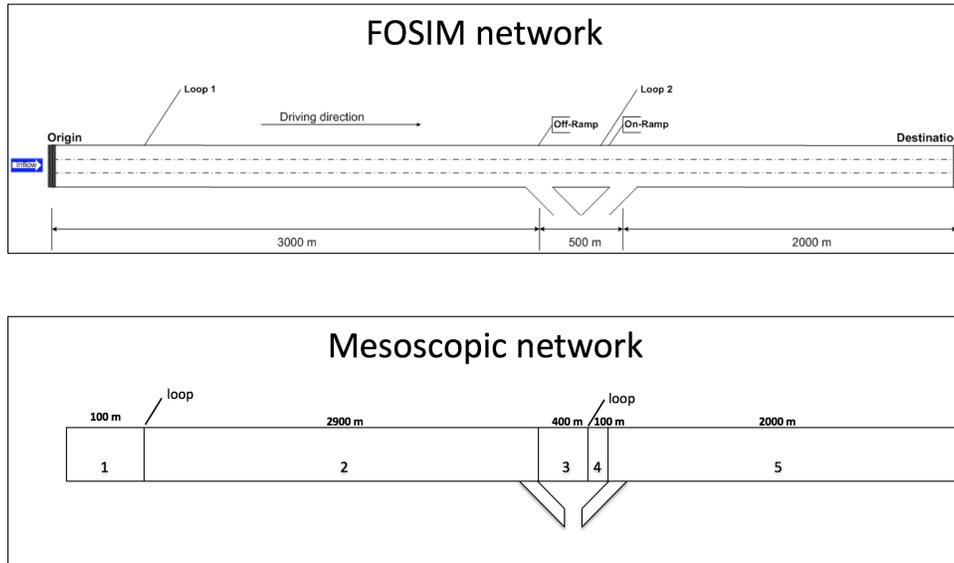


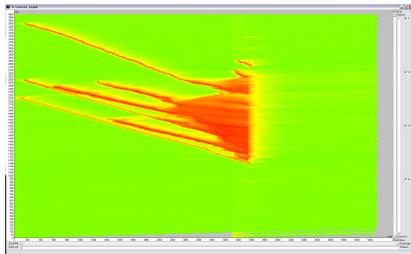
FIGURE 15: Networks

468 A demand-supply scenario has been built in such a way that a congestion
 469 is onset at the on-ramp. The model has been run twice : scenario 1 provides
 470 traffic with only passenger cars whereas scenario 2 considers a mixed traffic,
 471 with 90% cars, 10% trucks. The resulting time-space diagrams and travel
 472 times for the seed 1 are illustrated in Figure 16.

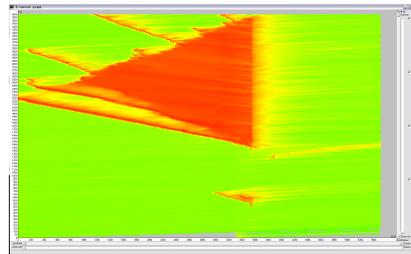
473 Based on FOSIM simulation results, Eulerian and Lagrangian observation
 474 models have been built. First, loop sensors have been located on the main
 475 road : loop 1 - 100m after the entrance of the network, and loop 2 - 100m
 476 upstream of on-ramp. Second, 10% equipped probe vehicles return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

478 5.2. Preparation of the traffic model

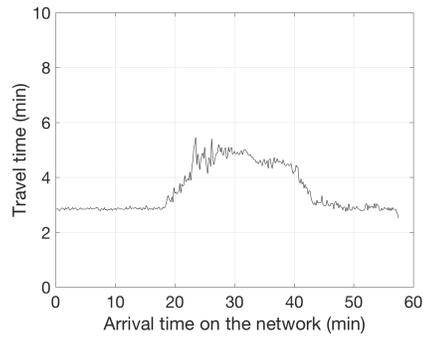
479 The traffic model is the mesoscopic LWR model applied on a network
 480 with 7 cells : five cells for the main road (numbered from 1 to 5), one cell 6
 481 for the off-ramp and one cell 7 for the on-ramp. Cell boundaries 1-2 and 3-4
 482 are located at loop sensor locations. Boundary conditions (demand-supply)
 483 are supposed to be known approximatively and parameters of the mesoscopic
 484 LWR have been set with the following default values : $u = 110$ km/h, $w = 18$
 485 km/h and $k_x = 150$ veh/km/lane. Results obtained from the underlying



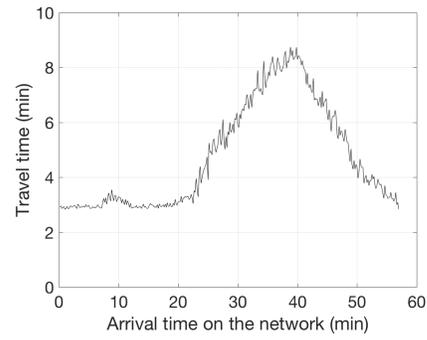
(a) time-space diagram



(b) time-space diagram



(c) travel times



(d) travel times

FIGURE 16: FOSIM observation models, seed 1 : cars only (*a* and *c*) and mixed traffic (*b* and *d*)

486 traffic model (without data assimilation) indicate that the corridor is free-
487 flowing, with travel times stabilized around 3 mins (175s).

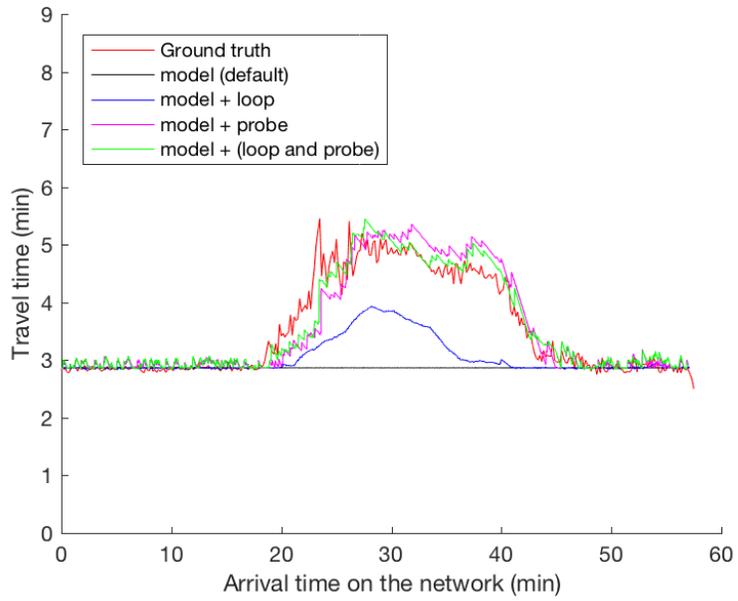
488 5.3. Results with different observation models

489 Three observation models have been tested : 'loop only', 'probe only',
490 and 'combined loop and probe'. Travel time estimation based on the three
491 observation models are illustrated in Figure 17, ground truth (in red) and
492 default model travel times (in black) are also displayed.

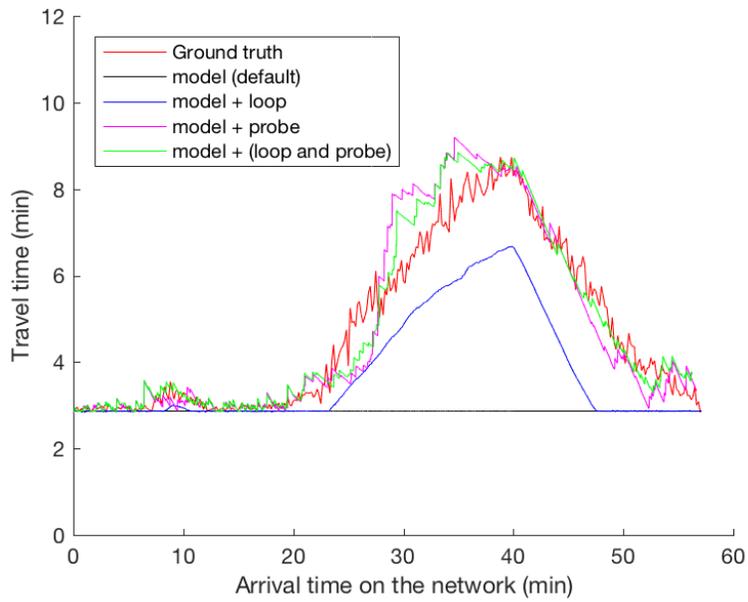
493 During the free-flowing period, travel times are properly estimated re-
494 gardless of observation models or traffic composition. However, significant
495 differences are observed when a congestion occurs. The results analysis only
496 focuses on the period $t = [20 - 45]min$ when the congestion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

497
498 *TSE based on the loop observation model* underestimates travel times
499 during the congestion period, regardless of traffic composition. This can be
500 caused by an underestimation of the upstream demand and/or an overestima-
501 tion of the supply. Loop 1 located at the entrance of the network is supposed
502 to update the demand according to the ground truth, so the overestimation
503 of the supply is the cause : loop 2 is located 100 meters upstream the head
504 of the congestion, which cannot detect immediately after its onset.

505
506 *TSE based on the probe observation model* presents a better performance.
507 However, it tends to overestimate travel times. It can be caused by a poor
508 prior estimation of demand, which skews the *n-index* estimation of probe
509 vehicle and leads to poor estimation of arrival/supply times. It can also be
510 caused by poorly calibrated traffic parameters in the traffic model and/or non
511 FIFO observations, which is confirmed in Figure 17(b) that shows the overes-
512 timation is enhanced for a mixed traffic (ranging from $t = [30 - 40]min$). By
513 analyzing FOSIM trajectories, it is observed that during congestion trucks
514 are stuck on the right-most lane (over congested) while most of the cars tra-
515 vel faster on left-most lanes. The FIFO assumption is not fulfilled and the
516 consequence on the performance of TSE can be explained as follows. When
517 a probe vehicle (for instance a truck) returns its position, arrival and supply
518 times are estimated in the (FIFO) mesoscopic model thereof. We conclude
519 that when the characteristics of probe vehicles are distant from the mean
520 traffic stream, this induces bias in traffic state estimation which tends to
521 overestimate travel times.



(a) Cars only, seed 1



(b) Mixed traffic, seed 1

FIGURE 17: Comparison of estimated travel times

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

526

527 Tables 3 and 4 provide three MoEs that have been calculated over the
 528 period $t = [20 - 45]min$: RMSE, MAPE and MPE. All the MoEs globally
 529 confirm the previous comments. In both scenarios, the estimation results with
 530 both data sources possess limited improvement compared to the probe-only
 531 cases. This can be explained by the fact that the demand input in the probe-
 532 only case well represents the actual demand so that the contribution from
 533 the loop data for flow correction is marginal. If the input demand contains
 534 noise, the improvement by assimilating additional loop data would be larger.

TABLE 3: Scenario with cars only - performance of the different observational models (10 replications)

	model	model + loop	model + probe	model + (loop and probe)
RMSE (s)	72	57	25	25
MAPE (%)	24	20	9	8
MPE (%)	-24	-19	0.65	0.43

TABLE 4: Scenario with cars and trucks - performance of the different observational models (10 replications)

	model	model + loop	model + probe	model + (loop and probe)
RMSE (s)	234	117	112	109
MAPE (%)	50	27	20	19
MPE (%)	-50	-27	8	14

535 5.4. Sensitivity to the penetration rate of probe vehicles

536 We expect that the performance of the data assimilation process can
 537 be improved when the amount of available data increase (n -, x - and t - fre-
 538 quency) and data are disaggregated. Here, we investigate the sensitivity of
 539 the proposed framework regarding the penetration rates of probe vehicles.
 540 The sensitivity has been tested on the scenario with cars only, with various
 541 penetration rates range from 0% to 10%. Figure 18 illustrates the sensiti-
 542 vity of the method regarding the penetration rates of probe vehicles (over

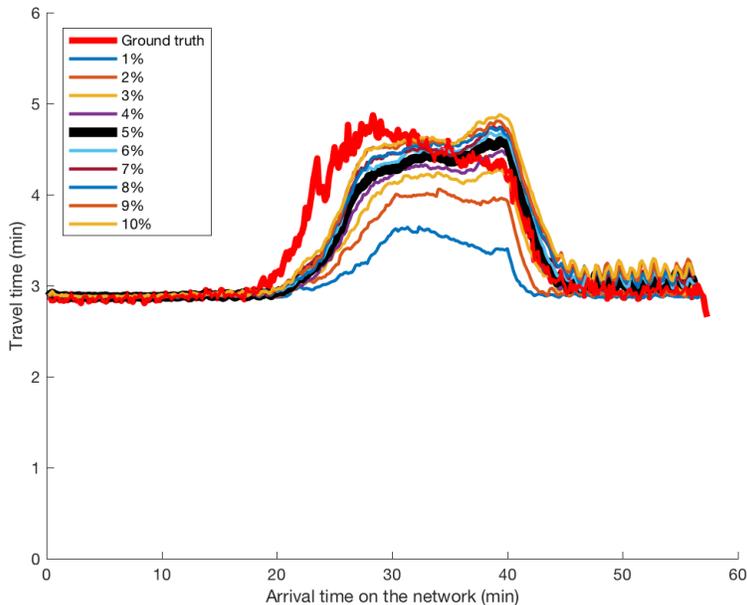


FIGURE 18: Travel time estimation for different probe penetration rates (10 replications)

543 10 replications). The estimation performance regarding MoEs is depicted in
 544 Figure 19.

545 The performance of the TSE is improved with only one percent of the
 546 observed probe vehicles. We also observe that the performance becomes mar-
 547 ginal as the percentage of probe vehicle exceeds 6%. It can be interpreted
 548 as follow : the gain for data assimilation is marginal and the remaining im-
 549 provement needs to be found elsewhere, for instance, with the three other
 550 aspects : First, the model relies on assumptions and limitations (triangu-
 551 lar fundamental diagram, FIFO traffic stream, etc.). Second, the definition
 552 of the demand on the network, which determines the traffic volumes that
 553 highly impacts the performance of the n -index estimation. And last but not
 554 least, the calibration of the model. We recall that the model has been poorly
 555 calibrated (default parameters) to emphasize the benefit of data assimilation
 556 procedure. Substantial gain is expected with a proper calibration procedure
 557 during the preparation of the model. We conclude that low penetration rates,
 558 around 5%, are sufficient for a realistic traffic state estimation.

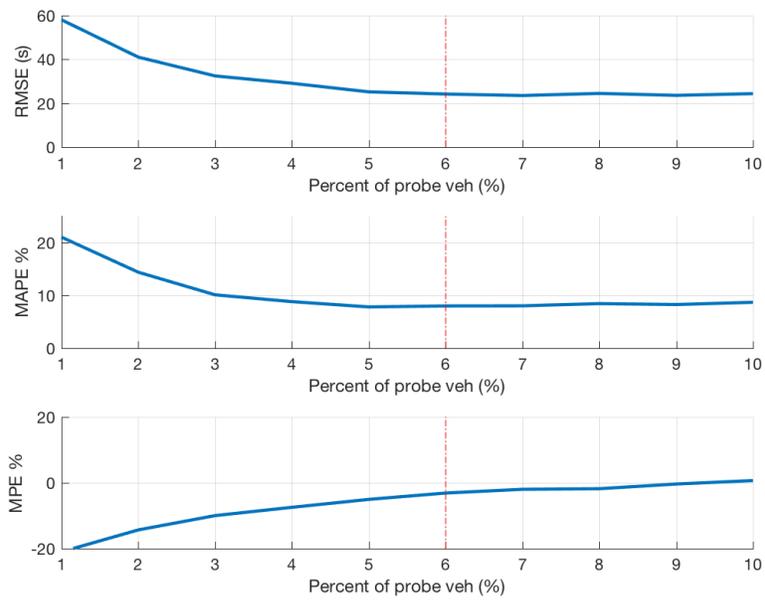


FIGURE 19: MoEs for different probe penetration rates (10 replications)

559 6. Discussion

560 Based on the previous results, we conclude that Eulerian observations can
561 update arrival and supply times, by adjusting the demand (flow) via adding or
562 deleting vehicles locally. However, since loop sensors are spatially fixed, they
563 only catch supply information as information propagates upstream. Hence,
564 travel time estimation might be inconsistent (see section 4.2, the case of the
565 downstream cell). It is therefore critical to locate loop sensors at the spots
566 of a jam and a bottleneck to provide accurate estimation.

567 In contrast, Lagrangian observations spread over the network in space
568 and time. Given a reasonable resolution (above 10%), they can update both
569 arrival times and supply times without any latency. However, two limitations
570 exist in this method. First, the *n-index* estimation relies on the assumption
571 of the FIFO condition, which is unrealistic. This might lead to poor esti-
572 mation of travel times when probe information deviates from the average
573 traffic conditions (see section 5.3). Second, the demand/flow at cell bounda-
574 ries cannot be adjusted. A prior estimation of the demand will improve the
575 performance of TSE with probe data (see section 4.2).

576 The combination of the two data sources compensates the limitations of
577 each other. The experiment results demonstrate that TSE with data combi-
578 nation outperforms the estimation with a single source.

579 7. Conclusion

580 7.1. Main findings

581 A TSE estimator based Eulerian observations combined with a mesosco-
582 pic LWR model has been proposed and validated in [12]. This paper com-
583 plements the methodology with Lagrangian observations. Now both Eulerian
584 and Lagrangian observations can be used for TSE in a unique framework.

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

591 Lagrangian observations provide a homogeneous coverage of the network
592 in time and space for a discrete set of (probe) vehicles. Probe vehicles allow
593 for an revision of demand-supply times at neighboring cell boundary of the

594 network. The update acts as a cell 'travel time regulator' that yields good
595 results under the condition that the demand on the network is known. Note
596 that the n -index values of probe vehicles are critical and essential in the
597 proposed approach. The calculation of this variable is application-specific
598 (under FIFO or non-FIFO condition) regarding estimation performance, and
599 it is subjective for further investigation.

600 In the TSE framework with data combination as proposed in section
601 3.3, Eulerian and Lagrangian observations become highly complementary.
602 Eulerian observations successfully update traffic states (especially the flow)
603 at loop locations of the network while Lagrangian observations successfully
604 update cell travel times along the network. The estimation outperforms the
605 scenarios considering loop and probe observations separately.

606 The methodology has been verified on the synthetic data derived from the
607 same underlying traffic flow model. Meanwhile, the proposed TSE framework
608 has been applied to a freeway corridor with a relaxed FIFO condition and
609 distributed driving behavior. The validity has been tested using the data
610 from a microscopic simulator, and the performance is satisfactory even for
611 low rate of probe vehicles around 5%. With increasing estimation accuracy
612 and computational efficiency, the proposed TSE framework will be beneficial
613 for decision support traffic management.

614 *7.2. Further research*

615 The robustness of data assimilation methodology is demonstrated in the
616 paper. However, we propose here some avenues to improve the present me-
617 thodology. On the model side, assumptions are to be relaxed to enhance the
618 model and its ability to reproduce well-documented traffic flow phenomenons
619 (non-FIFO condition, multi-class traffic, capacity drop). On the observation
620 side, the exponential growth data sources (e.g. bluetooth - mobile - infor-
621 mation from connected vehicles) will provide massive additional Lagrangian
622 and Eulerian information, which can be assimilated based on the proposed
623 framework providing their reliability and the proposition of adequate trans-
624 formation operators.

625 The data assimilation framework represents a solid base for on-line estima-
626 ting the reliability of both, again, traffic model and observed data. On the
627 model side, during the assimilation procedure, the discrepancy between the
628 background states from the model and the analysis states can be analyzed
629 for detecting incident on road networks and adjusting the model (demand,
630 parameters) accordingly. On the observed data side, the discrepancy between

631 the observed states and the analysis states can be help for estimating the ob-
632 servation reliability or detecting a problem with the data collection system.
633 And finally, we should recall that data assimilation has two main objectives :
634 to provide a continuous estimation of traffic states at the present time based
635 on discrete and aggregate observations ; and to propose the most possible
636 short term evolution of traffic states. The present paper is mainly focused on
637 the first item. Its capability to provide robust short term forecasts should be
638 explored, for instance, by implementing the proposed framework on a real
639 large-scale network.

640 **Acknowledgment**

641 The authors sincerely thank L. Leclercq (University of Lyon, ENTPE,
642 IFSTTAR, LICIT, UMR-T9401) for the fruitful discussions and assistance
643 on the methodology presented in the paper. We would also like to thank
644 anonymous reviewers for their valuable suggestions.

- 645 [1] Ansorge, R. : What does the entropy condition mean in traffic flow
646 theory? *Transportation Research Part B : Methodological* **24**(2), 133–
647 143 (1990)
- 648 [2] Aubin, J.P., Bayen, A.M., Saint-Pierre, P. : Dirichlet problems for some
649 hamilton-jacobi equations with inequality constraints. *SIAM Journal on*
650 *Control and Optimization* **47**(5), 2348–2380 (2008)
- 651 [3] Chu, L., Oh, S., Recker, W. : Adaptive kalman filter based freeway travel
652 time estimation. In : *Proceedings of the Transportation Research Board*
653 *84th Annual Meeting*. Washington D.C. (2005)
- 654 [4] Clairais, A., Duret, A., El Faouzi, N.E. : Calibration of the fundamen-
655 tal diagram based on loop and probe data. *Transportation Research*
656 *Record : Journal of the Transportation Research Board* (2560), 17–25
657 (2016)
- 658 [5] Claudel, C.G., Bayen, A.M. : Convex formulations of data assimilation
659 problems for a class of hamilton-jacobi equations. *SIAM Journal on*
660 *Control and Optimization* **49**(2), 383–402 (2011)
- 661 [6] Courant, R., Friedrichs, K., Lewy, H. : On the partial difference equa-
662 tions of mathematical physics. *IBM journal* **11**(2), 215–234 (1967)
- 663 [7] Daganzo, C.F. : A simple traffic analysis procedure. *Networks and Spa-*
664 *tial Economics* **1**(1-2), 77–101 (2001)
- 665 [8] Daganzo, C.F. : A variational formulation of kinematic waves : basic
666 theory and complex boundary conditions. *Transportation Research Part*
667 *B : Methodological* **39**(2), 187–196 (2005)
- 668 [9] Deng, W., Lei, H., Zhou, X. : Traffic state estimation and uncertainty
669 quantification based on heterogeneous data sources : A three detector
670 approach. *Transportation Research Part B : Methodological* **57**, 132–157
671 (2013)
- 672 [10] Dijkstra, T. : Fosim (freeway operations simulation) (2012). URL
673 <http://www.fosim.nl>
- 674 [11] Dimet, F.X.L., Talagrand, O. : Variational algorithms for analysis and
675 assimilation of meteorological observations : theoretical aspects. *Tellus*
676 *A* **38**(2), 97–110 (1986)

- 677 [12] Duret, A., Leclercq, L., El Faouzi, N.E. : Data assimilation using a meso-
678 scopic lighthill-whitham-richards model and loop detector data. Trans-
679 portation Research Record : Journal of the Transportation Research
680 Board **2560**, 26–35 (2016)
- 681 [13] Godunov, S.K. : A difference method for numerical calculation of discon-
682 tinuous solutions of the equations of hydrodynamics. *Matematicheskii*
683 *Sbornik* **47(89)**(3), 271–306 (1959)
- 684 [14] Herrera, J.C., Bayen, A.M. : Incorporation of lagrangian measurements
685 in freeway traffic state estimation. *Transportation Research Part B :
686 Methodological* **44**(4), 460–481 (2010)
- 687 [15] Kalnay, E. : Atmospheric modeling, data assimilation and predictability.
688 Cambridge university press (2003)
- 689 [16] Kordic, V. : Kalman Filter. Intech, Vukovar, Croatia (2010)
- 690 [17] Laval, J., Leclercq, L. : The hamilton-jacobi partial differential equation
691 and the three representations of traffic flow. *Transportation Research
692 Part B : Methodological* **52**, 17–30 (2013)
- 693 [18] Leclercq, L., Laval, J.A., Chevallier, E. : The lagrangian coordinates and
694 what it means for first order traffic flow models. In : R. Allsop, M. Bell,
695 B. Heydecker (eds.) *Proceedings of the 17th International Symposium
696 on Transportation and Traffic Theory*, pp. 735–753. Elsevier, London,
697 U.K. (2007)
- 698 [19] Lighthill, M., Whitham, G. : On kinematic waves ii : A theory of traffic
699 flow on long crowded roads. *Proceedings of Royal Society* **229A**(1178),
700 317–345 (1955)
- 701 [20] Nanthawichit, C., Nakatsuji, T., Suzuki, H. : Application of probe-
702 vehicle data for real-time traffic-state estimation and short-term travel-
703 time prediction on a freeway. *Transportation Research Record* **1855**,
704 49–59 (2003)
- 705 [21] Newell, G. : A simplified theory of kinematic waves in highway traffic,
706 part i : General theory. *Transportation Research Part B : Methodological*
707 **27B**(4), 281–287 (1993)

- 708 [22] Ngoduy, D. : Applicable filtering framework for online multiclass freeway
709 network estimation. *Physica A : Statistical Mechanics and its Applica-*
710 *tions* **387**(2/3), 599–616 (2008)
- 711 [23] Payne, H. : Models of freeway traffic and control. *Simulation Councils*
712 *Proceedings Series : Mathematical Models of Public Systems* **1**(1), 51–61
713 (1971)
- 714 [24] Richards, P. : Shock waves on the highway. *Operations Research* **4**(1),
715 42–51 (1956)
- 716 [25] Seo, T., Kusakabe, T. : Probe vehicle-based traffic state estimation meth-
717 od with spacing information and conservation law. *Transportation*
718 *Research Part C : Emerging Technologies* **59**, 391–403 (2015)
- 719 [26] Shannon, C.E. : Weaver ; w.(1949) : The mathematical theory of com-
720 munication. Urbana (1949)
- 721 [27] Talagrand, O. : Assimilation of observations, an introduction. *Journal-*
722 *Meteorological Society of Japan Series 2* **75**, 81–99 (1997)
- 723 [28] Velan, S., Florian, M. : A note on the entropy solutions of the hydro-
724 dynamic model of traffic flow. *Transportation Science* **36**(4), 435–446
725 (2002)
- 726 [29] Wang, Y., Papageorgiou, M. : Real-time freeway traffic state estimation
727 based on extended kalman filter : A general approach. *Transportation*
728 *Research Part B : Methodological* **39**(2), 141–167 (2005)
- 729 [30] Work, D., Tossavainen, O.P., Blandin, S., Bayen, A., Iwuchukwu, T.,
730 Tracton, K. : An ensemble kalman filtering approach to highway traffic
731 estimation using gps enabled mobile devices. In : *Proceedings of the*
732 *47th IEEE Conference on Decision and Control*, pp. 2141–2147. Cancun,
733 Mexico (2008)
- 734 [31] Yuan, Y., Duret, A., Van Lint, H. : Mesoscopic traffic state estimation
735 based on a variational formulation of the lwr model in lagrangian-space
736 coordinates and kalman filter. *Transportation Research Procedia* **10**,
737 82–92 (2015)

- 738 [32] Yuan, Y., Van Lint, J.W.C., Wilson, R.E., Van Wageningen-Kessels,
739 F.L.M., Hoogendoorn, S.P. : Real-time lagrangian traffic state estimator
740 for freeways. IEEE Transactions on Intelligent Transportation Systems
741 **13**(1), 59–70 (2012)