Transportation Research Part B 44 (2010) 460–481

Contents lists available at ScienceDirect

Transportation Research Part B

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

Incorporation of Lagrangian measurements in freeway traffic state estimation

Juan C. Herrera [a,*,1], Alexandre M. Bayen [b,2]

[a] Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile, Chile
[b] Systems Engineering, Department of Civil and Environmental Engineering, University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 15 May 2009
Received in revised form 26 October 2009
Accepted 26 October 2009

Keywords:
Lagrangian sensing
Traffic state estimation
GPS-enabled cell-phones
Probe data collection

ABSTRACT

Cell-phones equipped with a global positioning system (GPS) provide new opportunities for location-based services and traffic estimation. When traveling on-board vehicles, these phones can be used to accurately provide position and velocity of the vehicle as probe traffic sensors. This article presents a new technique to incorporate mobile probe measurements into highway traffic flow models, and compares it to a Kalman filtering approach. These two techniques are both used to reconstruct traffic density. The first technique modifies the Lighthill–Whitham–Richards partial differential equation (PDE) to incorporate a correction term which reduces the discrepancy between the measurements (from the probe vehicles) and the estimated state (from the model). This technique, called Newtonian relaxation, "nudges" the model to the measurements. The second technique is based on Kalman filtering and the framework of hybrid systems, which implements an observer equation into a linearized flow model. Both techniques assume the knowledge of the fundamental diagram and the conditions at both boundaries of the section of interest. The techniques are designed in a way in which does not require the knowledge of on- and off-ramp detector counts, which in practice are rarely available. The differences between both techniques are assessed in the context of the Next Generation Simulation program (NGSIM), which is used as a benchmark data set to compare both methods. They are finally tested with data from the *Mobile Century* experiment obtained from 100 Nokia N95 mobile phones on I-880 in California on February 8, 2008. The results are promising, showing that the proposed methods successfully incorporate the GPS data in the estimation of traffic.

1. Introduction

Traffic congestion in the US alone causes a 78 billion drain on its economy annually. In total, drivers lose 4.2 billion hours and waste 2.9 billion of fuel gallons per year because of traffic congestion (Schrank and Lomax, 2007). In order to improve the performance of the system, several traffic control strategies have been proposed and implemented over time. Coordination of traffic lights, ramp metering, dynamic speed limits, and information display through *changeable message signs* (CMS) are examples of such strategies.

Most of the highway flow control strategies require an accurate knowledge of the state of the system (for instance, vehicles accumulation or speed distribution on a highway segment). Traffic state estimation is a crucial part of any active control

* Corresponding author. Address: Casilla 306, Código 105, Santiago 22, Chile. Tel.: +56 2 354 4270; fax: +56 2 5530281.
E-mail addresses: jch@ing.puc.cl (J.C. Herrera), bayen@berkeley.edu (A.M. Bayen).
[1] Ph.D. student at UC Berkeley during the research presented in this article.
[2] Address: 642 Sutardja Dai Hall, Berkeley, CA 94720, United States. Tel.: +1 510 642 2468; fax: +1 510 643 5264.

0191-2615/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trb.2009.10.005

scheme on freeways, and usually requires the deployment of a significant monitoring infrastructure. The most common way to monitor traffic on freeways is the use of inductive loop detectors. These sensors are embedded in the pavement and collect data from vehicles as they pass over them. However, their high installation and maintenance cost prevents massive deployments, especially in developing countries.

Alternative ways to monitor traffic make use of information provided by individual vehicles. *Radio frequency identification* (RFID) transponders used to pay tolls (such as Fastrak in California or EZpass on the East Cost) are used to compute travel time between consecutive locations, which requires the installation of post readers along the road. This technology, however, is not universal and only provides point-to-point information. Probes vehicles reporting traffic conditions are also used for traffic monitoring purposes, but the penetration is not high enough to provide sufficient spatio-temporal coverage of the transportation network.

In the era of convergence of multimedia, communication and sensing platforms, GPS-equipped smart phones are becoming an essential medium for of location-based services. Given the high popularity of smart phones, this new sensing technology has the potential of providing a better coverage of the transportation network than current sensor technology does (and without public infrastructure costs associated with monitoring).

This new source of traffic data (speed measurements gathered from GPS-equipped phones on-board vehicles) creates a scientific challenge for traffic state estimation. Privacy issues and technology limitations (phone energy consumption and bandwidth usage) prevent the systematic collection of all data generated by the GPS. This means that the amount of data which can be used for traffic flow reconstruction is a subset of all data generated by each GPS.

This article presents a new algorithm to integrate GPS data in flow models, and compares it to an application of Kalman filtering for the specific problem of interest. Both methods assume the knowledge of the fundamental diagram and the conditions at both boundaries of the section of interest. The methods can work even when data is not available for the on- and off-ramps, which is almost always the case in practice. For this, the algorithms use data available at intermediate locations on the freeway, transmitted by the mobile devices.

The rest of the article is organized as follows: Section 2 provides some background, summarizing relevant work for traffic state estimation and introducing the method chosen to address the problem of interest. The methods are fully described in Section 3. Two data sets are used to test the methods. Section 4 describes both datasets and presents the main results. Finally, Section 5 states the main conclusions of the present study and proposes some future work directions.

2. Background

Traffic state estimation requires data, which may differ depending on the type of sensors used to collect it. Numerous freeways around the world, in particular in the US, are equipped with loop detector stations embedded in the pavement to collect traffic data. For each lane, these detectors aggregate information at a given sample time (usually 20 or 30 s). Vehicle counts, occupancy and speed are among the information that a detector can collect. These measurements provided by static sensors are traditionally referred to as *Eulerian* measurements, which means that the detector measures flow through a fixed control volume. On the other hand, *Lagrangian* sensors collect measurements of the system along a particle trajectory (motion of a car). RFID transponders, smart phones and GPS devices on-board vehicles providing position and/or velocity are examples of mobile (or Lagrangian) sensors.

Data can be used with different modeling approaches for traffic reconstruction. One class of models does not make use of traffic flow physics, and the estimation is based on statistics (current and historical data). Real-time traffic reports are sometimes based on this type of statistical models. These models have also been a common practice in studies that use smart phones as traffic sensors, in which one of the main goals has been speed or travel time estimation on a stretch of road (Sanwal and Walrand, 1995; Westerman et al., 1996; Ygnace et al., 2000; Bar-Gera, 2007; Krause et al., 2008). Note that the first four aforementioned studies use cell tower signal to obtain cell phone position, which is less accurate than GPS positioning. In Krause et al. (2008) the authors have investigated the use of machine learning techniques to reconstruct travel times on graphs based on sparse measurements collected from GPS devices embedded in smart phone and automobiles.

Treiber and Helbing (2002) propose a method that filters Eulerian data collected from loop detectors (density or velocity) to reconstruct traffic flow. In practice, density can be extracted from occupancy; therefore, density based approaches are possible. The filter is such that in free flow, traffic information propagates downstream, while in congestion it travels upstream. Even though this method takes into account the way in which information propagates in traffic streams, it does not make use of any flow model.

Another class of models is based on traffic flow physics. Lighthill and Whitham (1955) and Richards (1956) independently proposed a first order partial differential equation (referred to as the LWR PDE) to describe traffic evolution over time and space. The LWR PDE is a scalar hyperbolic conservation law, which relates changes in density over time to changes in flow over space. This law is based on the conservation of vehicles principle. Extensions of this model include a second equation accounting for the fact that vehicles do not accelerate/decelerate instantaneously, which are known as *second order models*, such as in Zhang (1998) and Aw and Rascle (2000). Numerical schemes, such as the Godunov scheme (Godunov, 1959), can be used to discretize these continuous models (both first and second order models). While the discussions with respect to first order and second order models led to specific conclusions on both sides, the method proposed applies to models of any

order. In the present case, we apply it to a first order model (the LWR model); extensions to second order models are straightforward, however, they are outside the scope of this article.

Daganzo (1994, 1995) proposed the *cell transmission model* (CTM), which is a discretization of the LWR PDE.[3] The CTM divides the highway in "cells" of length Δx and computes the state of the system (vehicle accumulation or density) every Δt units of time in each cell according to the conservation of vehicles principle. The CTM transforms the nonlinear flux function into a nonlinear discrete operator. Modifications to the CTM can be found in the literature. Muñoz et al. (2003) use a hybrid system framework to develop the *switching-mode model* (SMM), which combines discrete event dynamics estimation (mode identification) with nonlinear continuous dynamics state estimation (density estimation). Gomes and Horowitz (2006) modify the original merge rule of the CTM and propose the *asymmetric* CTM (ACTM).

For traffic state estimation, dynamic flow models of the system can be combined with data collected by sensors, a process known as data assimilation. However, there exist various techniques to perform data assimilation. Gazis and Knapp (1971), Szeto and Gazis (1972), and Sun et al. (2004) have used Eulerian measurements to perform data assimilation using Kalman filtering techniques. While Gazis and Knapp (1971) and Szeto and Gazis (1972) used the conservation of vehicles as model, Sun et al. (2004) uses the SMM. Chu et al. (2005) and Nanthawichit et al. (2003) have also performed data assimilation using Kalman filtering techniques, but with simulated Lagrangian data. Chu et al. (2005) use the conservation of vehicles equation and Kalman filtering techniques to combine point detection data and probe vehicle data into the travel time estimation, while a second order model is used in Nanthawichit et al. (2003) to perform data assimilation.

Data assimilation using Lagrangian data to estimate the state of a system is common in other fields such as meteorology and oceanography. In these fields, data assimilation methods range from simple, suboptimal techniques such as direct insertion, statistical correction, and statistical interpolation to more sophisticated, optimal algorithms such as inverse modeling, variational techniques and a family of methods based on Kalman filtering (Paniconi et al., 2003). The extrapolation of these techniques to transportation engineering problems appears to be very promising.

A known and simple method used in oceanography is Newtonian relaxation (or nudging) method (Anthes, 1974), which we use for the present study. The Newtonian relaxation method relaxes the dynamic model of the system towards the observations. To this end, a source term proportional to the difference between the predicted and observed state is included in the constitutive equation of the model (in the present case the LWR PDE).

Most of the data assimilation methods use boundary conditions at the boundary of the physical domain of interest. They are assumed to be known at some specific locations (usually at the boundaries of the computational domain). Thus, these methods make use of Eulerian data as well.

This article presents two methods for traffic state estimation that can handle both Eulerian and Lagrangian measurements, using a first order flow model. The contributions of this article specifically include:

- The development of a first order flow model which integrates a nudging term (to perform data assimilation) and its corresponding discretization using the Godunov scheme.
- An extension of the Kalman filtering based method used to estimate traffic state, following the work of Sun et al. (2003), which can incorporate Lagrangian observations.
- A Lagrangian data selection procedure to incorporate the measurement data into the discrete model.
- An implementation of both algorithms on an extensive dataset (the NGSIM data) for which ground truth is known.
- An implementation of both algorithms on the data set from the *Mobile Century* experiment, and a cross validation with PeMS data.[4]

3. Description of the approaches

The problem of interest for the present study is the incorporation of Lagrangian data into traffic flow models for freeways. The present section describes the new approach proposed to address this problem and the extension of the Kalman filtering approach specific to our problem. It first describes the flow model to be used and then presents the proposed methods.

3.1. Flow model approach

3.1.1. Continuous flow model

The continuous flow model proposed in the 1950s by Lighthill and Whitham (1955) and Richards (1956) describes the evolution of traffic on an infinite road. The LWR PDE relates density on the road and its flow:

$$\frac{\partial k(x,t)}{\partial t} + \frac{\partial}{\partial x}[q(k(x,t))] = 0 \tag{1}$$

The function $k(x, t)$ represents the vehicle density in vehicles per unit length and $q(k(x, t))$ is the flux function in vehicles per unit time (or the fundamental diagram, which is assumed to be triangular and time–space invariant in this work) at location x and time t.

[3] The CTM is a special case of the Godunov scheme when the fundamental relation between flow and density is assumed to be triangular.

[4] California Performance Measurement System: http://pems.eecs.berkeley.edu/.

The initial condition corresponds to the density along the road at the beginning of the period of analysis:

$$k(x,0) = k_0(x), \quad \text{for all } x \tag{2}$$

Given the dynamics of the traffic, the lack of accurate knowledge of initial conditions can be counter balanced by the "flush out" effect, i.e. for sufficiently large periods of time, the influence of the initial conditions becomes negligible.

For a finite road, the boundary conditions correspond to the density at the upstream and downstream ends of the section ($x = a$ and $x = b$, respectively), and they could be obtained from sensors at these locations, such as loop detectors. Let $k_a(t)$ and $k_b(t)$ be the *desired* density or the density measurements collected at time t by these sensors at boundaries a and b. Let $k(a, t)$ and $k(b, t)$ be the *modeled* density or the density computed with Eq. (1) at time t at the boundaries a and b. If boundary conditions are to be applied in the strong sense, the *desired* density and the *modeled density* should be equal all the time. That is, $k(a, t) = k_a(t)$ and $k(b, t) = k_b(t)$ for all t.

However, for a proper characterization of the solution to Eq. (1), weak boundary conditions are required, which implies that previous equalities between the *desired* and the *modeled* density do not hold all the time. Boundary conditions only apply (that is, the *modeled* density is made equal to the *desired* density) on the boundary of the section where characteristics are entering the computational domain. Otherwise, the *modeled* density will not necessarily be equal to the *desired* density.

The way in which characteristics move depends on the traffic conditions. In the free flow regime, characteristics travel downstream (i.e. in the same direction as the flow) while in congested regime they travel upstream, implying that characteristics enter the domain through the upstream (resp. downstream) boundary at location a (resp. b) if free flow (resp. congested) conditions prevail at this location. Therefore, the upstream boundary condition is relevant (and thus can be applied in the strong sense) only when a free flow condition is observed at that point. Otherwise, the boundary condition is irrelevant and conditions are dictated by downstream traffic. The opposite is true for the downstream boundary, in which the boundary condition is relevant only if congestion is observed at that point.

Mathematically, the weak boundary conditions for the specific case of the LWR PDE can be found in Bardos et al. (1979) for general hyperbolic conservation laws. In the present context, they can be written in simpler form as follows:

$$\left.\begin{array}{ll} \text{(i)} & k(a,t) = k_a(t) \quad \text{and} \quad q'(k_a(t)) \geq 0 \text{ or} \\ \text{(ii)} & q'(k(a,t)) \leq 0 \quad \text{and} \quad q'(k_a(t)) \leq 0 \text{ or} \\ \text{(iii)} & q'(k(a,t)) \leq 0 \quad \text{and} \quad q'(k_a(t)) \geq 0 \quad \text{and} \quad q(k(a,t)) \leq q(k_a(t)) \end{array}\right\} \text{for all } t \tag{3}$$

and

$$\left.\begin{array}{ll} \text{(i)} & k(b,t) = k_b(t) \quad \text{and} \quad q'(k_b(t)) \leq 0 \text{ or} \\ \text{(ii)} & q'(k(b,t)) \geq 0 \quad \text{and} \quad q'(k_b(t)) \geq 0 \text{ or} \\ \text{(iii)} & q'(k(b,t)) \geq 0 \quad \text{and} \quad q'(k_b(t)) \leq 0 \quad \text{and} \quad q(k(b,t)) \leq q(k_b(t)) \end{array}\right\} \text{for all } t \tag{4}$$

In Eqs. (3) and (4), $q'(k)$ is the slope of the flux function $q(k)$, defined as $q'(k) = \frac{dq}{dk}$. Similar conditions were derived by Daganzo (1995) from physical principles. Note that only (i) implies that the *modeled* density is equal to the *desired* density. In the other two cases, the *modeled* density is driven by the model and is not necessarily equal to the *desired* density.

The interpretation of the conditions in (3) is as follows (interpretation for condition (4) is analogous). One of the three alternatives is possible, based on the state of the system directly downstream $x = a$ (i.e. at $x = a + \varepsilon$, where ε is small):

- *Either (i)*: This means that it is possible to impose the *desired* density $k_a(t)$ upstream from the section of interest, therefore $k(a, t) = k_a(t)$. For example, a low flow situation in which the upstream inflow is dictated by the demand falls into this category.
- *Or (ii)*: Both the desired inflow and modeled inflow are congested flows, given by the negative slope of $q(\cdot)$. In this case, the solution (the boundary condition) is driven by the state, i.e. the inflow allowed by the current state of congestion.
- *Or (iii)*: The modeled flow is congested, but the desired inflow (demand) is not. However, the modeled flow value is greater than the desired flow value on the highway. In this case, the flow acceptable by the highway is smaller than the flow which the boundary condition seeks to push. Therefore, the boundary condition cannot drive the inflow.

While the physical reasoning between these three cases could probably be organized with a different Boolean logic, see in particular (Daganzo, 1995) for a physical analysis of this situation, formulas (3) and (4) above are mathematically required for the problem to be well-posed. In particular, it ensures existence and uniqueness of an *entropy solution* to this equation (Oleinik, 1957; Bardos et al., 1979) for a bounded domain, from which follows convergence of the Godunov discretization scheme, which is key to ensure the proper numerical solution to the problem. The first known instantiation of this Boolean type condition is due to Le Floch (1988), and more recently adapted for highway specific flux functions by Strub and Bayen (2006).

In practice, these boundary conditions are implemented using *ghost cells*. These cells correspond to the *input* and *output* cells as proposed by Daganzo (1994), and are presented in subsequent sections.

3.1.2. Discretization method

For implementation purposes, the LWR PDE needs to be discretized. To this end, the freeway section is divided into I cells (each one of length Δx distance units, and indexed by i starting form upstream). Time is divided into H time steps (each one of length Δt time units, and indexed by h). In order to meet the *Courant–Friedrichs–Lewy* (CFL) stability condition (LeVeque, 2002), which states that a vehicle traveling at the free flow speed v_f cannot traverse more than one cell in one time step, the condition $\Delta t \cdot v_f \leqslant \Delta x$ should be met. At every time step, the model estimates the density in each cell according to the following expression:

$$k_i^{h+1} = k_i^h - r\left(q_{i+1}^h - q_i^h\right) \quad i = 1, 2, \ldots, I, \quad \text{and} \quad h = 0, 1, \ldots, H-1 \tag{5}$$

The parameter r is the inverse of the speed needed to travel one cell in exactly one time step (i.e. $r = \Delta t/\Delta x$), while k_i^h is the discrete approximation of the density in cell i at time step h. The variable q_i^h is the flow into cell i between time h and $h+1$, and depends nonlinearly on the density of cells $i-1$ and i. It can be computed using the Godunov scheme as follows:

$$q_i^h = \begin{cases} q\left(k_i^h\right) & \text{if } k_c < k_i^h < k_{i-1}^h \\ q(k_c) & \text{if } k_i^h < k_c < k_{i-1}^h \\ q\left(k_{i-1}^h\right) & \text{if } k_i^h < k_{i-1}^h < k_c \\ \min\left\{q\left(k_{i-1}^h\right), q\left(k_i^h\right)\right\} & \text{if } k_{i-1}^h \leq k_i^h \end{cases} \tag{6}$$

In Eq. (6), k_c is the *critical density*, which corresponds to the point at which the flux reaches its only maximum (the flux function is concave).

As was explained earlier, weak boundary conditions are required for a proper characterization of the solution of the LWR PDE in (1). In the implementation, one ghost cell is inserted at each boundary of the section. Ghost cells are outside of the physical domain, and are virtual cells on which the Godunov update Eqs. (5) and (6) automatically apply weak boundary conditions (3) and (4), by construction of the Godunov scheme (LeVeque, 2002; Strub and Bayen, 2006). The ghost cells contain the boundary conditions and allow the first and last cells of the computational domain to be updated depending on the existing traffic conditions (free flow, congested, or a combination of the two). This approach was successfully implemented and tested in earlier work (Strub and Bayen, 2006) with traffic data collected from loop detectors. The equations for the ghost cells ($i = 0$ and $i = I+1$) and for the initial conditions ($h = 0$) are given by Eqs. (7) and (8), respectively:

$$k_0^h = \frac{1}{\Delta t}\int_{(h-1)\Delta t}^{h\Delta t} k_a(t)dt \quad \text{and} \quad k_{I+1}^h = \frac{1}{\Delta t}\int_{(h-1)\Delta t}^{h\Delta t} k_b(t)dt \quad h = 0, 1, \ldots, H-1 \tag{7}$$

$$k_i^0 = \frac{1}{\Delta x}\int_{x_{i-1}}^{x_i} k_0(x)dx \quad i = 1, 2, \ldots, I \tag{8}$$

3.2. Proposed methods

In order to perform traffic state estimation (vehicle accumulation or density, in this case) along the section of interest using Lagrangian data, specific data assimilation methods need to be developed. If accurate traffic data from all locations was available at all times, traffic state estimation would not be needed since the state can be directly inferred from the data. In practice, however, data is not always available from all the locations and is, in general, very sparse.

As explained in Section 2, different approaches can be used to obtain estimates for locations without measurements. The approaches adopted here use traffic flow models, based on vehicle accumulation (or density).

Two methods are proposed and discussed in the following sections, which rely on the following assumptions:

- GPS-enabled mobile phones traveling on the section of interest are sampled in time. They report their position and velocity at specific time intervals T.
- Boundary conditions are known. That is, the density at both ends of the section of interest is available. This data can be provided by loop detectors. However, as we shall see later, the availability of loop detector data at the boundaries is not a critical requirement.
- The fundamental diagram is assumed to be triangular and known.
- Information from intermediate ramps is not required. In fact, this constitutes one of the main features of the methods proposed since, in reality, information from ramps is rarely available.

3.2.1. Definitions: state, estimated, and observed variables

The *state variable* characterizes the state of a dynamical system. In the present case, it corresponds to vehicle density, and it is denoted by $k(x, t)$. By definition, it satisfies the physical model, which is the LWR PDE in (1) in the present case.

The *state estimate* (or estimated variable) is the result of the state estimation process, and it is distinguished from the model variable by the use of a *hat*. We denote $\hat{k}(x, t)$ the state estimate.

To perform the estimation, it is assumed that some variable is measured. In the present case, the available data consists of position and velocity measurements from GPS-equipped vehicles at different times. Therefore, a relationship between the velocity and the density is needed in order to relate the measured or observed variable and the estimated variable. This issue is addressed in the next paragraphs.

The observed position and velocity of an equipped vehicle labeled j at time t are denoted by $s_j^o(t)$ and $v_j^o(t)$, respectively (superscript o denotes *observation*). It is assumed that the observed velocity at the corresponding location and time is provided by the individual probe velocity measurement. Using $v(x,t)$ to denote the average velocity field on the freeway, we defined the observed velocity at time t and location $s_j^o(t)$ by:

$$v\left(s_j^o(t),t\right) = v_j^o(t) \tag{9}$$

This assumption may not be appropriate when dealing with multi-lane freeways, where different lanes might have different speeds. In these cases, this problem requires a specific treatment, which should use more sophisticated models, and are out of the scope of this work.[5] As a first approach, we treat the freeway as a single traffic stream.

The fundamental diagram relates the flow, the density and the velocity of the flow on a section of road. In particular, this relationship can be used to infer the density from the velocity. That is, using the fundamental diagram, $v^o\left(s_j^o(t),t\right)$ can be converted into the observed density at the point of measurement, denoted by $k^o\left(s_j^o(t),t\right)$. In reality, $k^o\left(s_j^o(t),t\right)$ is not the observed density but an estimate based on the speed measurement and the fundamental diagram. Since the fundamental diagram is only a model, and in reality flow-density points (or traffic states) do not necessarily lay on a line, this conversion is expected to introduce error in the observed density.

If a triangular fundamental diagram is used, a problem arises when the speed $v_j^o(t)$ is high. In fact, if the measured speed $v_j^o(t) \leq v_f$, where v_f is the free flow speed, different combinations of flow and density have the same free flow speed. Indeed, under free flow conditions it is not possible to observe the local density through speed measurements in the way described in the previous paragraph. For these cases, free flow conditions will be assumed and a "free flow density" value (denoted k^{FF}) will be used. The value for k^{FF} can change in time and space, and can be obtained from historical data. This seems reasonable considering that our main interest is to obtain accurate density estimates specifically when congestion arises.

Therefore, the observed density is given by:

$$k^o\left(s_j^o(t),t\right) = \begin{cases} \frac{w\cdot k_j}{v^o\left(s_j^o(t),t\right)+w} & \text{if } v^o\left(s_j^o(t),t\right) < v_f \\ k^{FF} & \text{otherwise} \end{cases} \tag{10}$$

Parameters k_j and w are the jam density and the slope of the right branch of the triangular fundamental diagram, respectively. The observed velocity $v^o\left(s_j^o(t),t\right)$ comes from Eq. (9).

3.2.2. Newtonian relaxation method

The Newtonian relaxation (or nudging) method is a simple heuristic method that has been used for data assimilation in the field of environmental fluid mechanics (Anthes, 1974). In oceanography, GPS-equipped drifters are used to estimate the velocity field of rivers, using shallow waters models. The extension of this technique to transportation engineering problems appears to be very promising, since GPS-equipped vehicles are similar to the drifters, and we aim to estimate the state of the freeway in terms of the vehicle density.

The Newtonian relaxation method relaxes the dynamic model of the system towards the observations. To this end, a source term called the nudging term, proportional to the difference between the estimated and observed state, is included in the constitutive equation of the model, which in the present case is the LWR PDE in (1):

$$\frac{\partial \hat{k}}{\partial t} + \frac{\partial q(\hat{k})}{\partial x} = -\sum_{j=1}^{J} \sum_{t_j^p \in \Omega_j^t} \lambda(x - s_j^o\left(t_j^p\right), t - t_j^p) \cdot \left[\hat{k}\left(s_j^o\left(t_j^p\right), t_j^p\right) - k^o\left(s_j\left(t_j^p\right), t_j^p\right)\right] \tag{11}$$

The summation over the index j in the RHS of Eq. (11) accounts for the J different vehicles equipped with Lagrangian sensors, while the second summation includes all the observations sent by each vehicle j before the current time t. The expression in Eq. (11) assumes that vehicle j sends observations at times $t_j^p \in \Omega_j^t$, where Ω_j^t represents the set of times until t at which measurements from vehicle j are performed and used for data assimilation (note that necessarily $t_j^p < t$ since only observations from the past are available at current time t).

The nudging factor $\lambda(\delta_x, \delta_t)$ represents the weight of each observation to be applied to the solution. This weight is expected to become negligible (and eventually to become zero) away from the measurement location and after the measurement time. For this reason, we have adopted here an expression for the nudging factor $\lambda(\delta_x, \delta_t)$ that takes this into account, and can be found in Ishikawa et al. (1996):

[5] In multi-lane freeways, more than one vehicle could send a measurement from the same location $s_j^o(t)$ at the same time t. In this case, the observed velocity is the average of all the measurements from location $s_j^o(t)$ at time t.

$$\lambda(\delta_x,\delta_t)=\begin{cases}\frac{1}{T_a}\exp\left(-\left(\frac{\delta_x}{X_{\text{nudge}}}\right)^2\right)\exp\left(-\frac{\delta_t}{T_d}\right) & \text{if } |\delta_x|\le X_{\text{nudge}} \text{ and } 0<\delta_t\leqslant T_d\\ 0 & \text{otherwise}\end{cases} \tag{12}$$

The factor dies out on a space and time scale of X_{nudge} and T_d, respectively. Therefore, close to where and when the observation is made, $\lambda(\delta_x, \delta_t)$ *nudges* the solution towards the observations. The parameter T_a has units of time and determines the strength of the nudging factor. A numerical example to show the effect of the nudging factor is provided in Appendix A.

In the present context, the nudging term "adds" or "removes" vehicles from the state of the flow model depending on whether the model underestimates or overestimates the number of vehicles on the freeway. In Section 3.2.3 we will discuss the implication of adding or removing vehicles on the conservation of vehicles principle.

3.2.2.1. Numerical implementation: discrete model. The discretization of the nudging term in the RHS of Eq. (11) needs to be added to Eq. (5). The final expression for the discretized model is as follows:

$$\hat{k}_i^{h+1}=\hat{k}_i^h-r(\hat{q}_{i+1}^h-\hat{q}_i^h)-\Delta t\sum_{j=1}^{J}\sum_{t_j^p\in\Omega_j^{h\Delta t}}\lambda\left(x_i-s_j^o\left(t_j^p\right),h\Delta t-t_j^p\right)\cdot\left[\hat{k}_{c_{jp}}^{m_{jp}}-k_{c_{jp}}^{o,m_{jp}}\right]$$
$$i=1,2,\ldots,I \quad \text{and} \quad h=0,1,\ldots,H-1 \tag{13}$$

The notation introduced in Eq. (13) is explained below:

- $\hat{q}_i^h$: estimated flow into cell i in time step h. This flow is obtained by application of (6) to the set of $\hat{k}_i^h$,
- x_i: location of the beginning of cell i, $x_i = x_0 + (i-1)\,\Delta x$, for $i = 1, 2, \ldots, I$, where x_0 is the beginning of the section of interest,
- c_{jp}: cell index corresponding to location $s_j^o\left(t_j^p\right)$, which is the location where the vehicle j is at time t_j^p when its pth observation is sent, $c_{jp}=\left\lceil\frac{s_j^o\left(t_j^p\right)-x_0}{\Delta x}\right\rceil$, where $s_j^o\left(t_j^p\right)>x_0$,
- m_{jp}: time step corresponding to location t_j^p, which is the time (in time steps units) when the pth observation from vehicle j occurs, $m_{jp}=\left\lceil\frac{t_j^p}{\Delta t}\right\rceil$.

The last term in the RHS of Eq. (13) is the discretization of the nudging factor times the difference between estimated and measured density (terms inside the square brackets). The nudging factor is given by the expression in (12).

3.2.3. Kalman filtering based method

Kalman filtering is a recursive method used to estimate the state of a discrete process governed by a linear stochastic dynamical system (Bar-Shalom and Li, 1993) in the presence of noisy measurements. The method assumes that the way in which the state of the system (density in this case) evolves is linear and known, and is referred to as the dynamics or state equation (which includes a process noise). Noisy measurements of the output of the system (i.e. observed density) are available, and the measurement or observation equation relates the output and the state of the system. Knowing the covariance of both the process and the measurement error, the method obtains the best estimate of the state of the system in the sense of the least squares.

Kalman filtering techniques have been proposed to perform traffic state estimation for loop detector data (Sun et al., 2004 and 2003). Given the nonlinearity of (6), the model in (5) is nonlinear as well. Therefore, conventional Kalman filtering can only be applied to linear subsets of this dynamics. This technique can be extended to cases in which data is provided by mobile sensors, which is described next.

3.2.3.1. State space representation. The state space representation of the system consists of two equations: the dynamics (or state) equation and the measurement (or observation) equation.

The dynamics equation describes how the state of the process (density) evolves over time and space. Since the constitutive equation in (5) is nonlinear, it needs to be linearized first. The hybrid system framework used in Muñoz et al. (2003) is adopted for this purpose. Discrete event dynamics estimation is performed to identify the traffic condition or mode of the section of interest. That is, the mode of each cell (i.e. free flow or congested) needs to be determined at the beginning of each time interval.

Previous studies (Muñoz et al., 2003; Sun et al., 2004) have proposed different ways to identify the mode on a short section of highway. For longer sections, the number of possible modes increases, adding complexity to the mode identification.

In the present study, state estimates (and indirectly Lagrangian measurements) are used to identify the mode. At the end of time interval h, density estimates for every cell are available. These estimates consider all the observations collected until time step h (inclusive) and are referred to as the *a posteriori* estimates at time step h. The *a posteriori* estimate at cell i at time step h is denoted by $\hat{k}_i^{+,h}$. The mode chosen for each cell will depend on the value of $\hat{k}_i^{+,h}$. In other words, if $\hat{k}_i^{+,h} > k_c$, cell i is

congested; otherwise, cell i is in free flow (k_c is the critical density). Note that the mode is being identified by using the state estimates and not by measuring the actual state of the system.

Once the mode for each cell has been identified, the flow into cell i between time step h and $h+1$, q_i^h, becomes linear in the densities in cells $i-1$ and i. Therefore, Eq. (5) can be written as follows:

$$\mathbf{k}_{h+1} = \mathbf{A}_h \cdot \mathbf{k}_h + \mathbf{B}_h \cdot \mathbf{u}_h + \mathbf{B}_h^{\mathrm{J}} \cdot k_J + \mathbf{B}_h^{\mathrm{Q}} \cdot q_{\max} + \mathbf{w}_h \tag{14}$$

Bold letters represent matrix or vector notation. The scalars k_J and $q_{\max}$ are the jam density and the maximum flow, respectively. The vector $\mathbf{u}_h$ is the input vector at time h, which includes the density at the boundaries of the domain, and $\mathbf{w}_h$ is the process error (caused for instance by the fact that not all the entry/exit counts are available). Eq. (14) is linear in the state $\mathbf{k}_h$, which is defined as $\mathbf{k}_h = \left[k_1^h \; k_2^h \cdots k_I^h\right]^T$. The matrices $\mathbf{A}_h$, $\mathbf{B}_h$, $\mathbf{B}_h^{\mathrm{J}}$ and $\mathbf{B}_h^{\mathrm{Q}}$ depend on the traffic conditions or mode of the section, which shows the importance of the mode identification step. The algebraic expression of these matrices for specific cases can be found in (Muñoz et al., 2003).

The measurement or observation equation projects the state vector into the measurements provided by Lagrangian sensors, $\mathbf{y}_h$, with the one predicted by the model:

$$\mathbf{y}_h = \mathbf{C}_h \cdot \mathbf{k}_h + \mathbf{v}_h \tag{15}$$

The vector $\mathbf{v}_h$ represents the measurement noise. The matrix $\mathbf{C}_h$ is time dependent, and its size and elements depend on the location of the measurements. It only contains zeros and ones, whose position in the matrix $\mathbf{C}_h$ depend on the locations at which the measurements are taken. The time dependency of $\mathbf{C}_h$ is a major challenge and is directly linked to the Lagrangian aspect of the measurements.

In summary, the dynamics or state equation and the measurement or observation equation of the system are given by Eqs. (14) and (15), respectively.

3.2.3.2. Kalman filtering. The following notation is used:

- $\hat{\mathbf{k}}_h$: *a priori* state estimate of $\mathbf{k}_h$, where $\hat{\mathbf{k}}_h = \left[\hat{k}_1^h \; \hat{k}_2^h \; \cdots \; \hat{k}_I^h\right]^T$,
- $\hat{\mathbf{k}}_h^+$: *a posteriori* state estimate of $\mathbf{k}_h$, where $\hat{\mathbf{k}}_h^+ = \left[\hat{k}_1^{+,h} \; \hat{k}_2^{+,h} \; \cdots \; \hat{k}_I^{+,h}\right]^T$,
- $\mathbf{P}_h$: *a priori* estimate error covariance, where $\mathbf{e}_h = \mathbf{k}_h - \hat{\mathbf{k}}_h$ is the *a priori* estimate error,
- $\mathbf{P}_h^+$: *a posteriori* estimate error covariance, where $\mathbf{e}_h^+ = \mathbf{k}_h - \hat{\mathbf{k}}_h^+$ is the *a posteriori* estimate error.

The difference between the *a priori* and *a posteriori* estimates at time step h is the fact that the *a priori* estimates do not take into account the observations collected at time step h, while the *a posteriori* estimates do. That is, the *a posteriori* estimate is an updated version of the *a priori* estimate.

Kalman filtering provides a set of recursive equations to estimate the vector state. The equations are as follows (Bar-Shalom and Li, 1993):

$$\hat{\mathbf{k}}_{h+1} = \mathbf{A}_h \cdot \hat{\mathbf{k}}_h^+ + \mathbf{B}_h \cdot \mathbf{u}_h + \mathbf{B}_h^{\mathrm{J}} \cdot k_J + \mathbf{B}_h^{\mathrm{Q}} \cdot q_{\max} \tag{16}$$

$$\mathbf{P}_{h+1} = \mathbf{A}_h \cdot \mathbf{P}_h^+ \cdot \mathbf{A}_h^T + \mathbf{Q} \tag{17}$$

$$\mathbf{F}_{h+1} = \mathbf{P}_{h+1} \cdot \mathbf{C}_{h+1}^T \left[\mathbf{C}_{h+1} \cdot \mathbf{P}_{h+1} \cdot \mathbf{C}_{h+1}^T + \mathbf{R}\right]^{-1} \tag{18}$$

$$\hat{\mathbf{k}}_{h+1}^+ = \hat{\mathbf{k}}_{h+1} + \mathbf{F}_{h+1} \cdot (\mathbf{y}_{h+1} - \mathbf{C}_{h+1}\hat{\mathbf{k}}_{h+1}) \tag{19}$$

$$\mathbf{P}_{h+1}^+ = (\mathbf{I} - \mathbf{F}_{h+1} \cdot \mathbf{C}_{h+1}) \cdot \mathbf{P}_{h+1} \tag{20}$$

Initial conditions $\hat{\mathbf{k}}_0$ and $\mathbf{P}_0$ are assumed to be known. $\mathbf{Q}$ and $\mathbf{R}$ are the covariance matrices of the process and measurement error, respectively. $\mathbf{F}_h$ is known as the Kalman gain at time step h. Note that there is a striking similarity between the two methods: the second term in Eq. (19) represents a non-physical source (correction) term, which modifies the *a priori* value of the state of the system, similar to the nudging term.

3.2.3.3. Implementation. The state vector contains the density in each cell. At the beginning of time step $h+1$, *a posteriori* estimates at time step h are available. The traffic conditions on the network at the end of time step h need to be identified in order to determine which set of matrices $\mathbf{A}_h$, $\mathbf{B}_h$, $\mathbf{B}_h^{\mathrm{J}}$ and $\mathbf{B}_h^{\mathrm{Q}}$ to use. The mode will be identified with the process outlined previously, which indirectly uses the Lagrangian observations collected. Note, however, that the mode identification is the most challenging task in the implementation of this method.

Once the mode has been identified at the beginning of time step $h+1$, Eqs. (16) and (17) are used to obtain the *a priori* density estimate and its covariance, respectively. At this point, Lagrangian data becomes available to the model, i.e. the observed local density at time $h+1$ will be known for some cells (the quantity and position of the Lagrangian sensors at $h+1$

will determine how many and for which cells the density is observed). With this information, the observed vector $\mathbf{y}_{h+1}$ and matrix $\mathbf{C}_{h+1}$ can be constructed. Then, the Kalman gain is computed using Eq. (18). Finally, the *a posteriori* density estimate and its covariance are obtained using Eqs. (19) and (20), respectively. In the event that no Lagrangian observation is available at time $h+1$, the matrix $\mathbf{C}_{h+1}$ is set equal to zero, which implies that the Kalman gain is also zero. In this case, the *a priori* and *a posteriori* density estimates are the same.

3.3. Comments on the proposed methods

Note first that in both methods, the observed density is computed using Eq. (10), which includes at least three error sources. The first source of error has to do with Eq. (9), when the velocity of an individual vehicle is assumed to correspond to the velocity at that given location. The second source of error is related to the fact that the fundamental diagram is not exact. Thus, the velocity-to-density conversion, as expressed in Eq. (10), introduces error. This error may not be negligible for free flow because of the approximation made for these cases.[6] The third source of error corresponds to the measurement error in the velocity $v_j^o(t)$, which is expected to be small given the accuracy of GPS.

Both methods are conceptually similar. They add or remove vehicles depending on the difference between an estimated density and the observed density computed using GPS data. For this, they used a so-called "observer equation", which is derived from the flow model, but includes modifications which integrate the measurements. The methods differ, however, in the way this difference between measurement and estimate is used. The Kalman filtering based method assumes that an observation obtained from cell i at time step h only corrects or updates the density on the corresponding cell and time step. The flow model propagates the effect in time and space. On the other hand, the Newtonian relaxation method uses the observation to directly affect the density of neighboring cells and for future times. Therefore, the effect of each observation is propagated in time and space directly through the nudging term, but also through the flow model. This may be useful when the number of observations is low, because if cell i is congested at time step h (according to the observation), neighboring locations are expected to also be congested for a certain period of time (congestion takes time to vanish).

Some tradeoffs between the two methods include the following:

- *Modeling complexity*: The Newtonian relaxation method is trivial to develop on almost any flow model, since it only consists in adding a source term weighed properly by the nudging factor. For Kalman filtering, the difficulty consists in finding the proper linearization, which in the present case cannot be obtained by linearizing the dynamics directly, but by identifying the proper modes (hence the hybrid system approach), which is a research topic in itself.
- *Implementation difficulty*: The benefit of the Newtonian relaxation method is such that any solver used for forward simulations (of the model) can be directly used and modified to include the nudging term, a feature which has been extensively used in other fields, in particular in oceanography. For Kalman filtering, additional update equations (outlined earlier) need to be added.
- *Parameter tuning*: The nudging term requires tuning three parameters, which are usually chosen based on physical considerations, while the Kalman filtering approach requires assumptions on the covariance matrices.
- *Computational cost*: Both methods have roughly the same cost (i.e. the cost of a forward computation), with additional computations required for the Kalman filter because of the observer equation.
- *Convergence of the estimator*: While the convergence of the Kalman filter can be assessed theoretically (Bar-Shalom and Li, 1993), Newtonian relaxation is a heuristic technique for which no convergence results is known to the authors.
- *Convergence of the model*: By definition of the Godunov scheme used, the discrete model used in Newtonian relaxation without the nudging term can be shown to converge to the solution of the LWR PDE. An extension of this property can be used to develop a convergence result for a PDE which includes a continuous nudging term. No such counterpart is known to the authors for the Kalman filter.

As was briefly explained before, mass (vehicle) conservation is not satisfied by the methods proposed in this article. This is a common feature of estimation techniques, so the next paragraphs explain the considerations which have to be taken into account when doing estimation, and how our methods fit in them.

Traditionally, estimation makes the assumption that the physics of a phenomenon are governed by a constitutive model (Bar-Shalom and Li, 1993). The equations for such models can be of various nature, in particular partial differential equations like Eq. (1) or difference equations like in Eq. (5). Because of modeling inaccuracy (models are never perfect), and because of measurement noise (measurements are never noise free), empirically measured data almost never satisfies a model perfectly, i.e. plugging the data in a model will violate the model. In the present case, the model is represented by Eq. (1) with corresponding continuous state variable $k(x,t)$, and the corresponding discrete model by Eq. (5), with discrete state variable k_i^h. With empirically measured data, $k(x,t)$ would never satisfy Eq. (1) exactly, and k_i^h would never satisfy Eq. (5) exactly.

The field of estimation has produced a large number of techniques which are based on so-called observer equations, which are traditionally denoted by hats (therefore $\hat{k}(x,t)$ and $\hat{k}_i^h$), which integrate measurements into constitutive equations,

[6] Note that if density estimates are used to obtain travel times (by computing the velocity), the value used for k^{FF} would not affect the travel time estimates significantly. Therefore, travel times estimates might be more accurate than density estimates under free flow conditions.

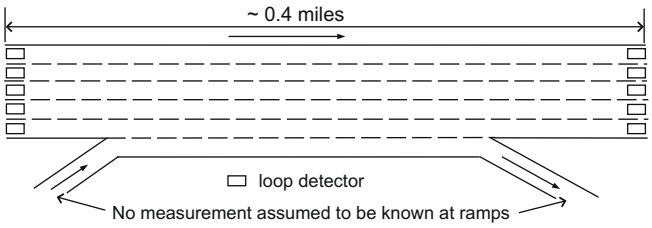

Fig. 1. Highway US101 S used for the NGSIM dataset.

but usually do not satisfy modeling assumptions anymore (i.e. in the present case, $\hat{k}(x,t)$ and $\hat{k}_i^h$ will not satisfy mass – or vehicle – conservation). The field of estimation and data assimilation is based on constructing evolution equations for these estimator states ($\hat{k}(x,t)$ and $\hat{k}_i^h$ in the present case), which are shown to minimize discrepancy error between measurements and model estimates. For example, Kalman filtering equations add a corrective term to the model equations (therefore violate the model), and provide a least square estimator of the state of the system.

Both the Newtonian relaxation method and the Kalman filtering based method add or remove vehicles from a cell depending on the relative values of the observed and estimated density. If the counts were known exactly at every entry and exit point of the network, this addition/removal of vehicles would violate the conservation of vehicles. In practice however, on- and off-ramps are rarely equipped with loop detectors (in addition to the fact that specifically to the case of the highway, the observed quantities are never perfect because of measurement errors). In particular, loss or gain of vehicles in between loop detectors is a well known problem in traffic engineering.[7] The present methods can thus be used to incorporate Lagrangian data in place of this missing loop detector data. The methods thus conveniently bypass the modeling of networks: for mainlines with on- and off-ramps only (no major intersection between highways), they replace the merge–diverge junctions by Lagrangian data incorporation.

When no loop detector data is available at ramps, numerical simulations are simply underdetermined, because of the lack of inflow and outflow information. When data from ramps is available, because of measurement errors, the data might be inconsistent with the model, in particular vehicle conservation. The methods presented steers the state of the model locally (in x and t) towards the Lagrangian measurements, which is a way to reestablish a value of the state closer to the actual state of the system wherever such measurements are available. Therefore, it compensates for the lack of inflow and outflow information or the corresponding inaccuracy.

Because of the considerations discussed above, the conservation of vehicles that incorporates the on- and off-ramps is replaced by the nudging term (in the Newtonian relaxation method) or the correction term (in the Kalman filtering based method), which perform the required local adjustments based on measurements.

4. Assessment of the methods

The Newtonian relaxation method (NR) and the Kalman filtering method (KF) are implemented on two datasets. In the first one, we have full knowledge of all positions and speeds of all vehicles during the entire experiment, which enables an extensive validation of the method against "ground truth". In the second one, we use GPS data obtained from smart phones from a subset of drivers and thus demonstrate the applicability of the methods to this novel way of gathering traffic data.

4.1. Implementation with NGSIM data

Traffic data from the *Next Generation Simulation* (NGSIM) project[8] is used to evaluate the proposed approach. NGSIM data has been extracted from video, which provides ground truth trajectories for all vehicles. The data consist of the trajectories of all vehicles entering a stretch of US Highway 101S in Los Angeles, CA, during 45 min (from 7:50am to 8:35am on June 15, 2005). The site is approximately 0.4 miles in length, with five mainline lanes (see Fig. 1). An auxiliary lane exists between the on-ramp and the off-ramp. Transition from free flow to congestion happens during the first 10–12 min of the data set (i.e. part of the section was in the free flow mode while the rest was in the congested mode).

For implementation purposes, the five mainline lanes were considered. The full section (1920 ft) has been divided into $I = 16$ cells of $\Delta x = 120$ ft each, and the total simulation time (2610 s) has been divided into $H = 2175$ time steps of $\Delta t = 1.2$ s each. Parameters of the fundamental diagram – maximum flow ($q_{\max}$), jam density (k_j), critical density (k_c), free flow speed (v_f) and wave speed (w) – were extracted from the PeMS system.[9]

[7] There are algorithms to correct for the measurement errors. For instance, miscounts arising during congestion can be corrected once free flow conditions are restored. See the appendix in (Chung et al. 2007) for more details.

[8] http://ngsim.camsys.com/.

[9] Assuming a triangular shape of the fundamental diagram: $q_{\max} = 2040$ vphpl, $k_j = 205$ vpmpl, $k_c = 30$ vpmpl, $v_f = 68$ mph, and $w = -11.7$ mph.

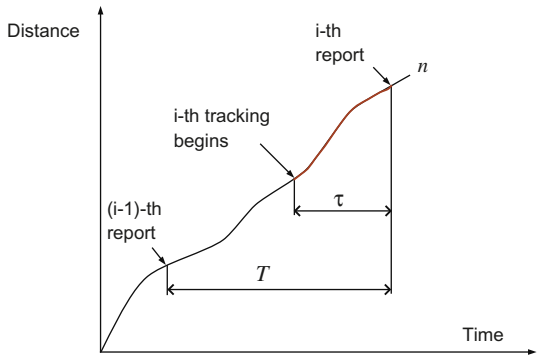

Fig. 2. Schematic of the sampling strategy on equipped vehicle *n*.

Table 1
Twelve scenarios investigated in the NGSIM data set.

Case	*P* (%)	*T* (s)	# Of Lagrangian measurements per mile-lane per minute
1	2	150	1
2	5	150	3
3	10	150	6
4	15	150	10
5	20	150	13
6	25	150	16
7	2	10	9
8	5	10	22
9	10	10	42
10	15	10	62
11	20	10	80
12	25	10	97

The data is processed in order to emulate loop detector data at both boundaries of the section. Since all vehicle trajectories are known, this can easily be done without usual measurement error associated with loop detectors. These emulated detectors will provide boundary conditions as defined in Section 3.2. Ramps counts could be inferred in the same manner, but they are assumed to be unknown, a realistic assumption given the lack of available infrastructure for most freeways in North America.

4.1.1. Scenarios investigated

4.1.1.1. Penetration rate and sampling strategies. The implementation of the algorithms selects a subset of the NGSIM data and treats it as Lagrangian data used for the assimilation. Twelve different scenarios were investigated to account for different penetration rates and sampling strategies. The penetration rate P is the proportion of trajectories that are chosen (randomly) as equipped vehicles. These equipped vehicles report their position and speed every time interval T. The reported speed is the average speed over the last τ seconds (Fig. 2 sketches this sampling strategy). The values chosen for P, T, and τ determine the total number of Lagrangian measurements created for each case investigated. Five different values of P and two values of T were investigated (τ was assumed 6 s in all scenarios). Table 1 shows the scenarios investigated and the average number of Lagrangian measurements per mile-lane per minute for each one of them.

The travel time for the section of interest is around 2 min under congested conditions. Thus, scenarios with $T = 150$ s assume that each equipped vehicle sends only one report while it is traveling the section. Almost continuous tracking for equipped vehicles is assumed for scenarios 7–12 ($T = 10$ s and $\tau = 6$ s). Because of privacy issues, it is not clear if such a sampling strategy would be socially acceptable for a commercial product, a problem addressed in (Hoh et al., 2008). This issue is still open and generates ongoing debates. They were investigated to evaluate the traffic reconstruction potential of the proposed method.

4.1.1.2. Parameters selection. The nudging factor $\lambda(\delta_x, \delta_t)$ depends on three parameters. Nudging parameters are typically set heuristically based on physical considerations. Parameters X_{nudge} and T_d determines how far in space and time, respectively, an observation influence the solution. For instance, if congestion is observed at location x at time t, it is expected that congestion also exists at locations in $[x - X_{\text{nudge}}, x + X_{\text{nudge}}]$ between times $[t, t + T_d]$. For the present case, the values for X_{nudge} and T_d are 180 ft and 15 s, respectively. The parameter T_a can be seen as a gain that determines the strength of the nudging term. Therefore, it has to be tuned as a control parameter of the model. For the present study, values of 10, 20 and 30 s have provided good results for the scenarios investigated.

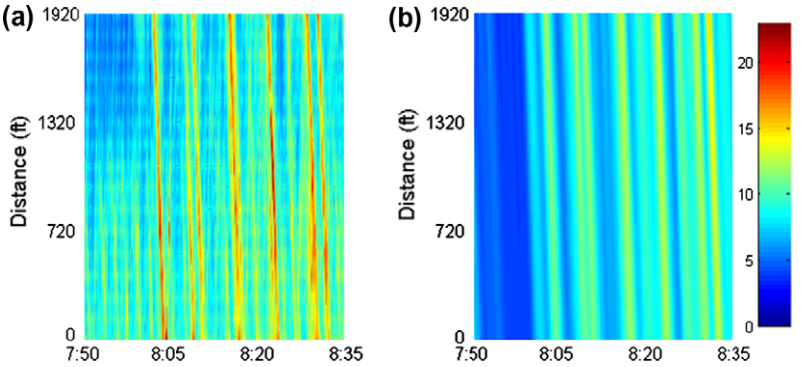

Fig. 3. Vehicle accumulation per cell for (a) ground truth and (b) EDO case.

4.1.2. Results

The 12 scenarios presented in Table 1 were implemented following the methods described in Section 3. For comparison purposes, we use a scenario which only incorporates information from boundary detectors, which we will refer to as *Eulerian data only* (EDO). This scenario was implemented according to the numerical scheme presented earlier, and does not make use of the ramp counts for the estimation.

The *Root Mean Square Error*[10] (RMSE) of the accumulation of vehicles (i.e. number of vehicles per cell per time step) is used as a metric for accuracy of the method. The true accumulation of vehicles can easily be computed since all vehicle trajectories are known. The RMSE for the EDO case is 2.6.

Fig. 3 shows the ground truth and estimated evolution over time (horizontal axis) and space (vertical axis) of the accumulation of vehicles for the EDO case. Colors close to red correspond to high accumulation of vehicle (i.e. congestion). Fig. 4 shows the same information for scenarios 1, 3, and 9 using the NR (left) and the KF (right). It can be seen that all the estimates are able to capture the main shockwaves traversing the section of interest. The intensity of the shockwave, however, is not equally captured by different scenarios.

Figs. 3 and 4 suggest that the main difference among the estimations occurs during the first 10–12 min (before 8:00am), when some waves emanate from intermediate locations. In fact, the fundamental difference between the EDO scenario and scenarios with Lagrangian measurements – in terms of the RMSE – happens during this period of time. The added value of Lagrangian measurements is thus clear, as it enables the methods to capture phenomena otherwise not detectable with loop detectors only. For the remaining period of time (until 8:30am), the scenarios with Lagrangian measurements do not show a significant improvement when compared with the EDO. Fig. 5 shows the evolution of the actual total accumulation of vehicles on the entire section over time and its corresponding estimate with EDO and scenario 5 for NR and KF. This graph confirms that the main difference occurs during the first 10–12 min.

It is expected that for longer sections between detectors, with more intermediate ramps and probably more waves emanating from intermediate locations, the difference between EDO and scenarios with Lagrangian measurements would be larger.

Table 2 shows the RMSE for each scenario investigated and its corresponding *Percentage of Improvement* (PoI) when compared with the EDO case. The PoI for scenario i is computed as $\text{PoI}_i = \frac{\text{RMSE}_{\text{EDO}} - \text{RMSE}_i}{\text{RMSE}_{\text{EDO}}} \cdot 100$. Note that the RMSE for each scenario corresponds to the average over 20 different realizations.[11] Fig. 6a shows the same information as a function of the total number of Lagrangian measurements used.

Fig. 6a suggests an increasing performance of the algorithms as the number of observations increases. The performance reaches a saturation level for large number of observations (more than 40 observations per mile-lane per minute in this case) at about 22% for the NR and 31% for the KF. The solid line shows the performance for scenarios 1–6 (one observation per equipped vehicle), while the dotted line corresponds to scenarios 7–12 (almost continuous tracking). For the NR, the lines in Fig. 6a also suggest that for similar numbers of observations (scenarios 4 and 7), having more vehicles sending observations less frequently seems to produce more accurate results than having few vehicles being continuously tracked. This gap is not observed in the KF implementation.

Ideally, the *observed* density computed using the observed velocity and the fundamental diagram (Section 3.2.1.) should perfectly match the ground truth density. As expected, it is not the case in practice. Fig. 7 shows the true density – computed from vehicle trajectories – vs. the *observed* density computed according to Section 3.2.1. for scenarios 3 and 9 (the trend is very similar for the other scenarios). This figure reveals the differences between the true and the observed density, suggesting that: (i) the speed reported (and used to compute the observed density) is not representative of the actual state on the

[10] The RMSE is defined as: $\text{RMSE} = \sqrt{\frac{\sum_h (\hat{z}_h - z_h)^2}{H}}$, where $\hat{z}_h$ and z_h are the estimator and its actual value at time step h, respectively, and H is the total number of observations.

[11] For the same penetration rate P and sampling strategy (P, τ), different realizations consider different vehicles as equipped vehicles and different times when measurements are sent.

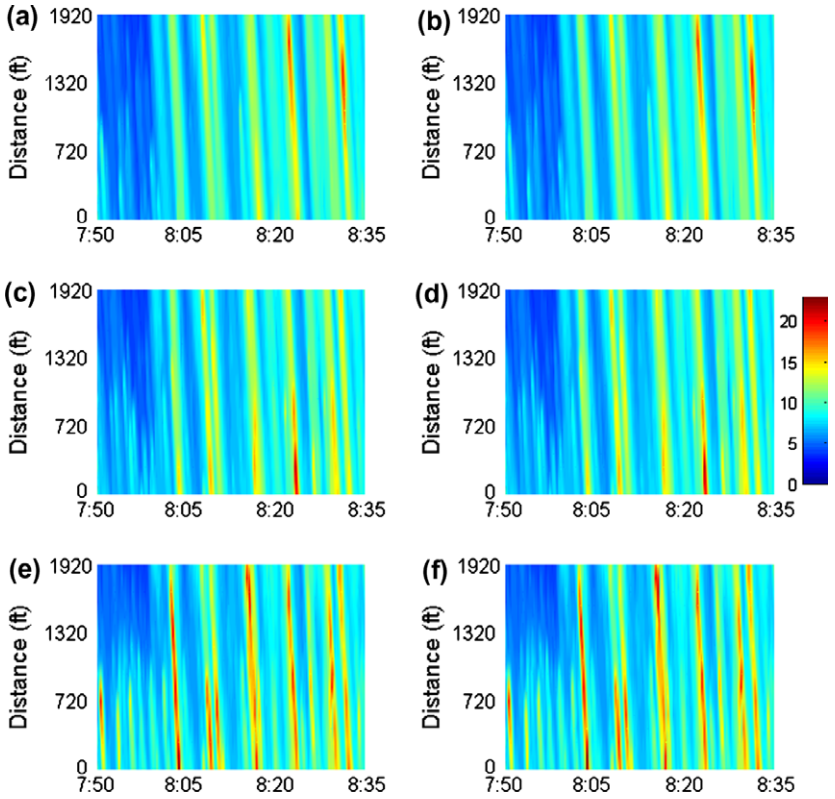

Fig. 4. Vehicle accumulation (vehicles per cell) estimated using Newtonian relaxation method (left) and Kalman filtering techniques (right) for scenarios 1 (top), 3 (middle), and 9 (bottom).

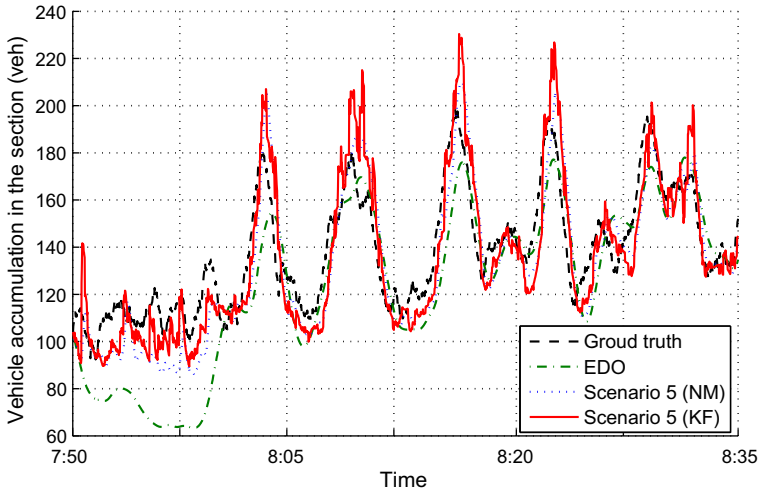

Fig. 5. Total vehicle accumulation on the entire section.

cell, and/or (ii) in congestion, a given density can be achieved at different velocities (or same velocity yields different densities). We believe that the main source of error is related with (ii). Indeed, it is well known that the congested branch of the fundamental diagram is not a line of points but a cloud of points for non-stationary traffic (Cassidy, 1998).

To determine how the error in the observed density affects the performance of the methods, all scenarios were investigated assuming that the density computed from the Lagrangian observations was the actual one. That is, the observations are perfect (error-free). The improvement of the RMSE with respect to the EDO case is shown in Fig. 6b. As expected, the accuracy of the estimates improves, achieving up to 40% (NR) and 50% (KF) of total improvement compared with the EDO case.

Table 2
RMSE for each scenario and its improvement with respect to the EDO case.

	Newtonian relaxation		Kalman filtering	
Scenario	RMSE	Improvement (%)	RMSE	Improvement (%)
1	2.44	6.1	2.33	10.1
2	2.35	9.8	2.18	16.2
3	2.21	14.9	2.04	21.6
4	2.16	16.7	1.98	23.7
5	2.14	17.7	1.94	25.1
6	2.11	18.9	1.90	27.0
7	2.27	12.8	2.00	23.2
8	2.18	16.2	1.89	27.4
9	2.06	20.6	1.81	30.3
10	2.02	22.3	1.80	30.7
11	2.03	22.0	1.78	31.3
12	1.98	23.8	1.78	31.6

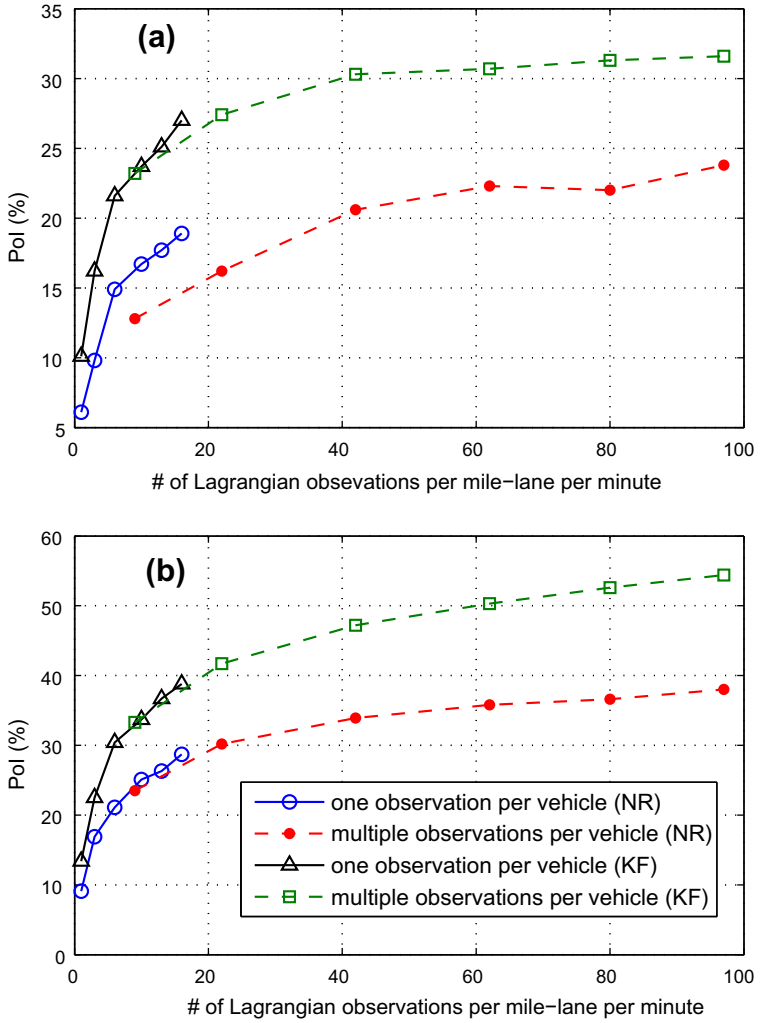

Fig. 6. Percentage of Improvement (PoI) in the RMSE as the number of Lagrangian measurements changes (Sce: scenario): (a) computing observed density using the fundamental diagram (Section 3.2.1) and (b) using the actual density computed from vehicle trajectories as the observed one. Note that both graphs are at different scales.

The gap between the solid and the dotted line observed in Fig. 6a for the NR is not observed in part b of the figure, suggesting that when the Lagrangian observations report the correct density, the sampling strategy does not really matter.

Statistical approaches such as the ones mentioned in Section 2 can be used to compute the velocity field. Using the fundamental diagram, the velocity field can be converted into the density field. This process is useful to compare the results

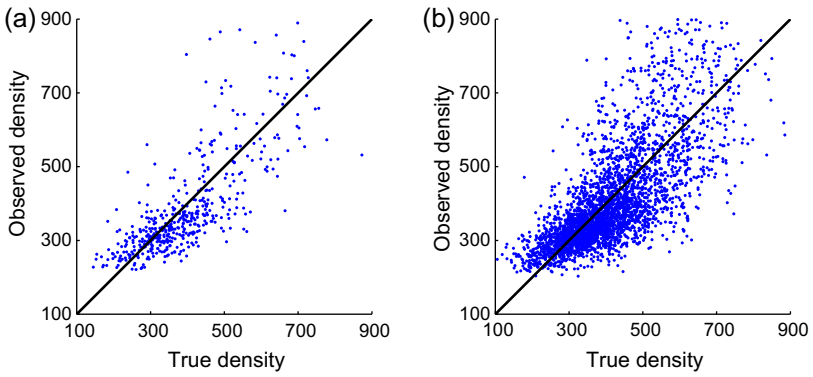

Fig. 7. True (computed from vehicle trajectories) vs. observed density (computed using the fundamental diagram as described in Section 3.2.1) for scenario 3 (left) and 9 (right).

obtained with the Newtonian relaxation method and the results obtained with a simple statistical approach – in this case, the method based on the approach proposed by Sanwal and Walrand (1995). In terms of the RMSE, only scenarios 11 and 12 show an improvement compared to the EDO case (3.4% and 5.7%, respectively). The rest of the scenarios yield a RMSE higher than the one obtained for the EDO case. That is, for this case, a method based on a traffic flow model produce more accurate estimates of the vehicle accumulation than a simple statistical method. This is not surprising since the statistical method is meant to estimate velocity and not vehicle accumulation.

The main difference between the estimates from the EDO scenario and from scenarios using Lagrangian measurements was observed during the periods in which shockwaves emanate from intermediate points. This result is insightful regarding the benefits of using Lagrangian measurements. By collecting data from individual vehicles at different times and locations, we were able to capture the shockwave generated in the middle of the section (between detectors). Therefore, in the presence of Lagrangian data, inter detector spacing could be increased. Unfortunately, the NGSIM section is too short to test this statement. The next section describes and provides result of a field deployment conducted to obtain Lagrangian measurements from GPS-enabled smart phones on a longer section of highway.

4.2. Field experiment: Mobile Century

Nicknamed the *Mobile Century* experiment, a field experiment was carried out on February 8, 2008, involving 100 vehicles carrying GPS-enabled Nokia N95 phones. These vehicles drove repeated loops of 6–10 miles in length continuously for 9 h on freeway I-880 near Union City, CA (see Fig. 8) in order to achieve up to a 5% penetration rate of the total volume of traffic on the highway during the experiment. 165 UC Berkeley drivers were hired for this experiment.[12]

This section of highway was selected specifically for its interesting traffic properties, which include alternating periods of free flow and congestion throughout the day. In particular, the northbound direction (NB) presents a recurrent and severe bottleneck between Tennyson Rd. and CA92 during the afternoon. On the day of the experiment there was an accident during the morning that activated a non-recurrent bottleneck at this same location. The section is also well covered with existing loop detector stations feeding into the PeMS system, which was used to assess the quality of the estimates.

The N95 Nokia phones used for the study stored their position and velocity data every 3 s, which after processing the data enables the reconstruction of every equipped vehicle trajectory.

For operational reasons, lanes 1–4 (lane 1 being the leftmost lane) on the section between Decoto Rd. and Winton Ave. in the NB direction were considered. The period of time analyzed starts at 10am and ends at 6pm. The 6.5 miles were divided into $I = 44$ cells of $\Delta x = 780$ ft each, and the 8 h period of meaningful data was divided into $H = 3600$ time steps of $\Delta t = 8$ s each. As in the NGSIM case, parameters of the fundamental diagram were extracted from PeMS.[13]

4.2.1. Different scenarios

The N95 Nokia phones stored position and velocity logs every 3 s. The penetration rate P is given by the number of smart phones on the road. During the day, this rate ranged between 2% and 5% of the total flow. Using this data, we changed the values of T to construct different scenarios following the method outlined in the previous section ($\tau = 2$ for all three scenarios, where τ corresponds to the number of logs over which the average speed is computed). Table 3 shows the scenarios investigated in this case and the total number of Lagrangian measurements per lane-mile per hour that each one produces.

The parameter choice for X_{nudge} and T_d is the same as for the NGSIM case (180 ft and 15 s, respectively). The values of T_a chosen for this experiment are 10 s for the first two scenarios and 30 s for the third one.

[12] A full description of the experiment is visible in video format at http://traffic.berkeley.edu.

[13] Assuming a triangular shape (and using the same notation as before): $q_{\max} = 2275$ vphpl, $k_j = 152$ vpmpl, $k_c = 35$ vpmpl, $v_f = 65$ mph, and $w = -19.4$ mph.

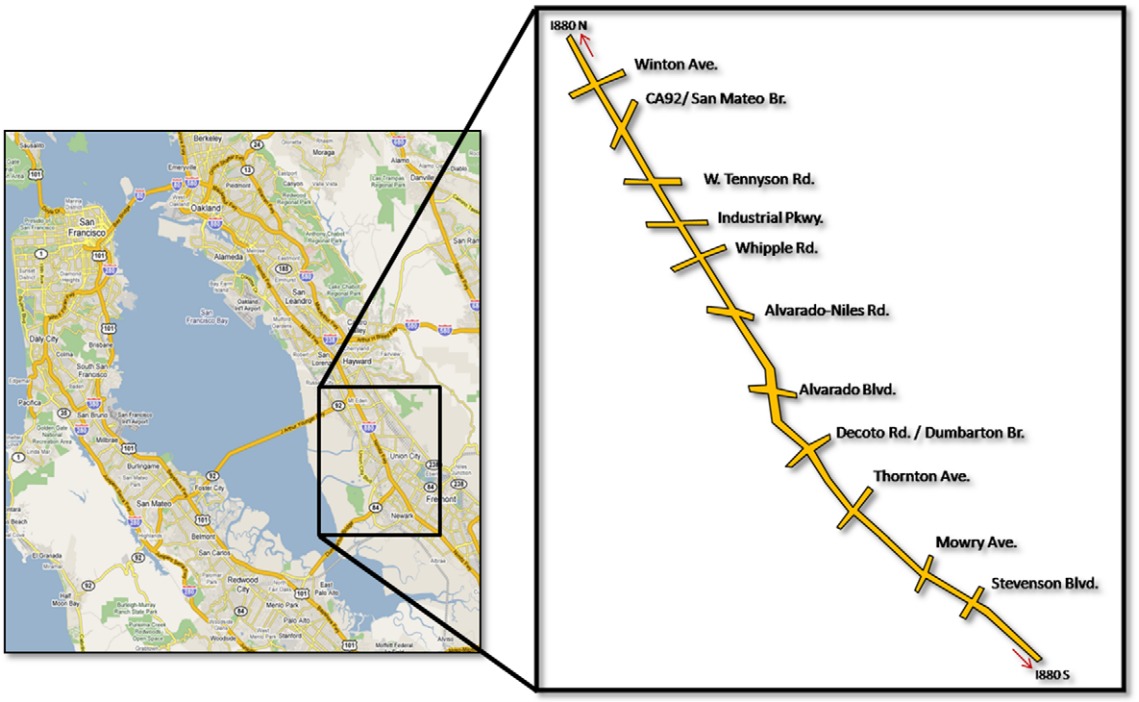

Fig. 8. Deployment section on highway I-880, Ca, for the *Mobile Century* experiment.

Table 3
Scenarios investigated using data from field experiment.

Case	T (s)	# Of Lagrangian measurements per mile-lane per minute
1	300	0.2
2	60	1.1
3	10	5.1

4.2.2. Results

Unlike for the NGSIM data, in this case we only have access to trajectories of the equipped subset of vehicles. Therefore, trajectories and vehicle accumulation ground truth are not known. To evaluate the results we use information collected by 17 loop detector stations installed along the section of interest using the PeMS system.

Fig. 9 shows the density field produced by PeMS. The density field collected by PeMS captures the morning accident at postmile 26 (between Tennyson Rd. and CA92). It also shows the spatial and temporal extent of the queue that was formed upstream. The situation at this location does not fully recover, and the bottleneck remains active until the evening. A short wave propagates upstream for about 2 miles at 12:30pm. At 2pm congestion starts, and by 3pm all the section between postmile 21 (Decoto Rd.) and 26 exhibit some serious level of congestion. Note that the severity of congestion around postmile 24 (between Whipple Rd. and Industrial Blvd.) is higher than for some other locations, suggesting the presence of a second active bottleneck in series at this postmile.

Fig. 10 shows estimated density for each scenario in Table 3 using NR (left) and KF (right). Recall that each scenario only makes use of two detector stations (at both boundaries of the section) in addition to the Lagrangian data provided by the cell phones. The main congested pattern observed from detectors is well captured in the three scenarios investigated and for the two methods. All the scenarios captured the accident at postmile 26 at 10:30am, even though not all the equipped vehicles were on the highway by then (vehicles were released between 10 and 10:30am). The location, intensity and duration of the incident are well replicated in each scenario. The short wave at 12:30pm is also observed on the three scenarios, but with different duration and intensity. The main congestion that starts at 2pm can be seen in the three scenarios as well. The more severe congestion observed during the afternoon from detectors around postmile 24 is also observed from Lagrangian data.

Fig. 10 also shows that KF results appear *noisier* than the NR results, especially during the free flow conditions. This can be seen by the yellow *points* observed in the graphs before PM 23 and 3pm, which correspond to observations with high density. In these cases the effect is visually magnified by the fact that traffic is freely flowing in the surroundings. This difference between NR and KF can be explained by the way in which each method uses each observation. While NR distributes the effect

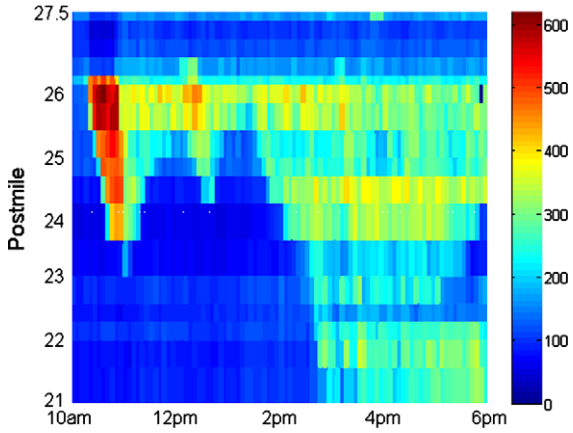

Fig. 9. Density field (in vpm) using 17 loop detector stations deployed in the section of interest (obtained through PeMS).

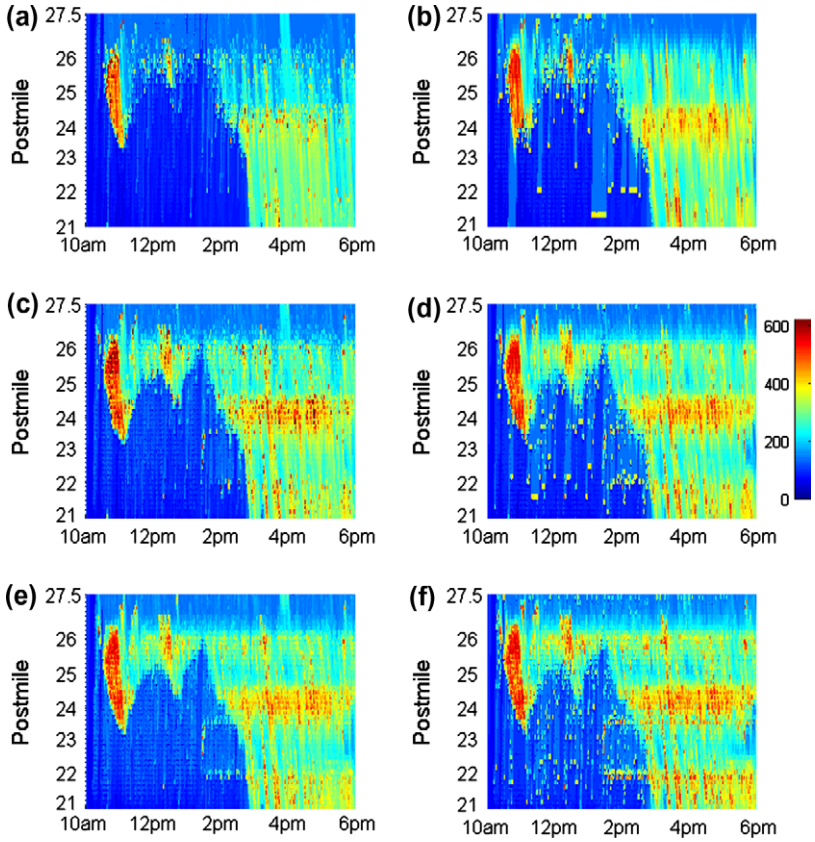

Fig. 10. Density field (in vpm) using the Newtonian relaxation method (left) and the Kalman filtering techniques (right) for scenario 1 (top), 2 (middle), and 3 (bottom). For each scenario, the boundary data is provided by loop detectors.

of the observation among several cells, KF directly affects only the observed cell. This means that the addition of vehicles in the KF is concentrated mainly in the observed cell.

In summary, there is a good agreement between the density field provided by PeMS and the reconstruction using Lagrangian data. As expected, the agreement with PeMS data improves with the number of Lagrangian observations.

The value chosen for k^{FF} in Eq. (10) was arbitrarily set to 25 vpmpl. Different values of k^{FF} change the estimates but not significantly. The main difference observed for different values of k^{FF} occurs during the transition between free flow and congested periods. Cases with higher values of k^{FF} tend to congestion more easily than lower values.

Note that the estimates shown in Fig. 10 make use of only two detector stations located at both boundaries of the section. However, in the absence of detector data at the boundaries of the section, the results are still very accurate. For instance, we can assume that the Eulerian boundary data is also provided by equipped vehicles. To this end, the concept of *virtual trip lines* (VTL) introduced in Hoh et al. (2008) is useful. A VTL can be thought of a virtual loop detector that records speed of equipped vehicles as they cross it. In our case, every 5 min a VTL at each boundary of the section of interest computes the average speed from GPS measurements from different vehicles. Obviously, the more equipped vehicles per interval, the more representative the estimate of the speed should be. Knowing the velocity, the fundamental diagram can be used to compute the density at the boundaries, which is the input to the model. When VTLs are used to provide boundary data the results do not change significantly. There is still a good agreement between the density field from PeMS and the estimates for each scenario. In fact, when boundary data is assumed to be totally unknown and some reasonable but constant value is assumed, the qualitative results still agree well with those obtained from PeMS.

Finally, the flows between consecutive cells can also be used for comparison. Fig. 11 shows a flow comparison at three existing loop detector locations. Each graph shows the flow measured by the corresponding loop detector and the three flows yielded by each scenario using the NR. Given that the correction terms in both methods are in part correcting the error in the flows between cells and the measurement error in loop detectors, a perfect agreement between the estimated flows and the flows measured by loop detectors cannot be expected. Despite the previous reason, and the fact that only two detectors located 6.5 miles from each other are being used in addition to the Lagrangian data, the estimated flows are in reasonable agreement with the flows measured by the detectors.

4.2.2.1. Quantitative analysis. Ground truth density is not known. The density field obtained using the 17 loop detector stations shown in Fig. 9 is available. Measurements from these stations, however, have errors. For this reason, a quantitative analysis cannot use loop detector measurements as ground truth. Moreover, loop detectors measure the occupancy, which is converted into density using some approximations.

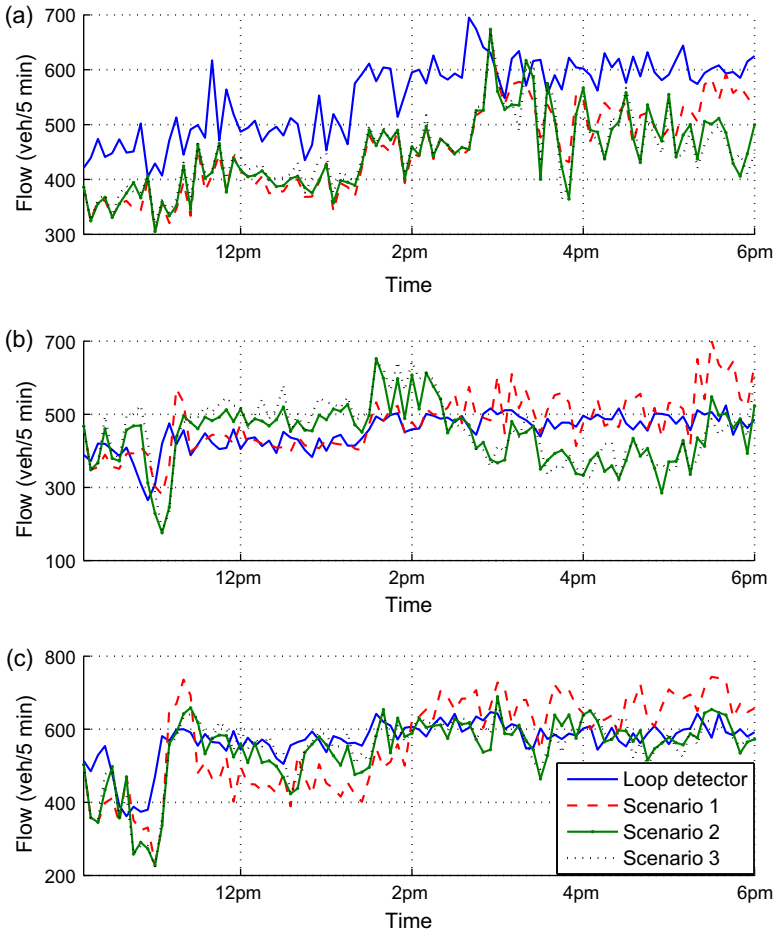

Fig. 11. Flow comparison at postmiles (a) 21.3 (detector 1), (b) 24 (detector 7), and (c) 25.2 (detector 10).

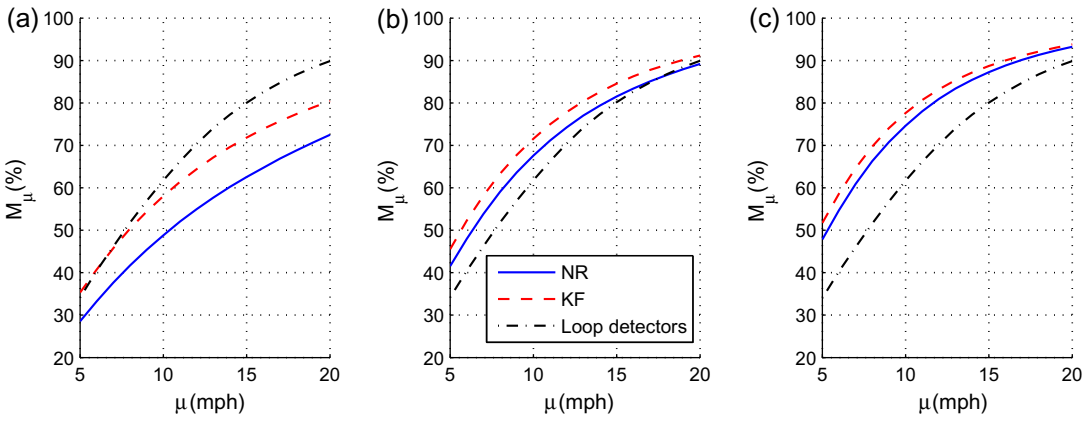

Fig. 12. Results of the quantitative analysis for the Newtonian relaxation method and the Kalman filtering technique for scenario 1 (left), 2 (center), and 3 (right). Results obtained using loop detector data are also included for comparison.

The procedure adopted to assess quantitatively the results consists of using 20% of the trajectories obtained in the *Mobile Century* experiment to provide ground truth velocity. The ground truth velocity is computed as the average of all the velocity logs in the corresponding cell at the corresponding time step. Only 20% of the trajectories, however, are not enough to provide with a velocity measurement for every cell at every time step. Therefore, velocity ground truth is only known for some cells at some time steps, and the comparison can only be performed at these cells and time steps (which represents around 10% of all the cell-time step points in the grid).

Ground truth is in terms of velocity, but the output of the methods is an estimate of the density field. The estimated velocity field is obtained using the fundamental diagram and by reversing the process described in Section 3.2.1. Note that for traffic state estimation using Lagrangian data, the other 80% of the trajectories are used in the estimation, reducing by 20% the number of observations for all the scenarios.

The metric used to assess the performance of the methods corresponds to the proportion of the estimates with less than μ mph of absolute difference with respect to the corresponding ground truth measurement, denoted by M_μ. The value for μ ranges between 5 and 20 mph.

Fig. 12 shows the results. Note that values presented are the average obtained from 20 realizations, where each realization considers a different set of vehicles in the 80% used for the estimation (and thus in the 20% used for the validation as well). For comparison purposes, results obtained using loop detector velocity measurements have also been included. For scenarios 2 and 3, velocity fields obtained with the methods proposed have a better agreement with the velocity field obtained from the validation set than the velocity field obtained using loop detector measurements from 17 stations. For high values of μ, all the three methods tend to converge in these cases. When the number of observation is very low (like in scenario 1), the estimated velocity field using measurements from 17 loop detector stations is preferred. The results also suggest that Kalman filtering techniques outperform the Newtonian relaxation method in all the scenarios investigated. Interestingly, both methods in the presence of Lagrangian data achieve similar (or even better) performance levels than the level achieved by a good coverage of loop detector stations.

5. Final comments

5.1. Conclusion from this study

Two methods to integrate Lagrangian measurements into a traffic flow model to perform traffic state estimation have been proposed. Lagrangian measurements used in the estimation consist in position and velocity readings from GPS-enabled smart phones traveling on the road. The methods were tested on two data sets: the NGSIM data set for validation purposes, and the Mobile Century data set because of its relevance for this application.

The main conclusions of this work are:

- The Newtonian relaxation method (NR) and the Kalman filtering techniques (KF) appropriately incorporate the Lagrangian measurements into flow models for traffic state estimation purposes. This is true even for a low number of observations per mile-lane per minute. This statement validates the value of both methods for traffic state estimation in the presence of Lagrangian data.
- Lagrangian measurements from vehicles (i.e. speed) are converted into density using the fundamental diagram, which introduces some error. However, despite this error the methods proposed produces accurate estimates. When the actual density is obtained from Lagrangian measurements, the accuracy of the results improves.

- In both methods, the accuracy of the final estimate grows with the number of Lagrangian measurements. The improvement in the estimates reaches a saturation level for high number of Lagrangian measurements. Marginal benefits are obtained from new observations when the number of measurements becomes too large.
- In both implementations, KF yields better results that the NR. In the implementation using the NGSIM dataset, the measure of performance was the percentage of improvement of the root mean square error of the density field. This index for the KF is approximately 1.5 times the index of the NR. In the implementation with the *Mobile Century* data, the measure of performance was the proportion of the estimates with less than certain level of absolute difference with respect to the corresponding ground truth measurement. The index for the KF is greater than the index for the NR, although the difference decreases when more Lagrangian observations are available. The differences observed between both methods can be explained by the fact that KF provides an optimal estimate in the sense of the least squares, while the estimates obtained with the NR are not necessarily optimal. Note again that in the KF algorithm, a major assumption is made when the mode is identified directly from the hybrid system model. This assumption might explain why for these datasets the method performs better (in fact, in the case of the NGSIM dataset, no mode identification is required since the traffic is congested all the time).
- Even though having more observations is better, results from the NGSIM analysis, and using the Newtonian relaxation method, suggest that it is better to have less frequent sampling on more vehicles, than frequent sampling on few vehicles. This is explained by a better spatial and temporal coverage of the first strategy.
- The proposed method makes use of density measured at upstream and downstream boundary. This information is assumed to be provided by loop detectors. However, when no detector data is available, GPS-enabled smart phones can provide this information without affecting significantly the performance of the method. Therefore, in the presence of Lagrangian data, loop detector stations are not essential for traffic state estimation.
- Both approaches are particularly useful when no information from ramps is available, and therefore current Eulerian data based models cannot handle conservation of vehicles.

The article thus demonstrates the added value of Lagrangian information in the system. It also highlights the potential of this type of data by itself (self sufficient for traffic reconstruction at high enough penetration rates). It finally underlines the tremendous potential of this type of data assimilation: even without fixed detectors, this data has the potential of supplying transportation agencies with loop detector quality (and potentially above) data, at no additional public infrastructure cost.

5.2. Future work

The nudging factor determines the desired effect of each Lagrangian measurement over time and space on the estimates of density. Potentially, a nudging factor that takes into account the way in which traffic information flows in different conditions could yield better results. For instance, an observation reporting congestion at location x_c at time t_c should influence locations upstream of x_c at times $t > t_c$. Based on this, different functional form for the nudging factor can be investigated to determine its influence in the results.

The present study proposes the use of Kalman filtering techniques to incorporate Lagrangian measurements in the traffic state estimation, following the work of Sun et al. (2004). This approach requires the identification of the mode at each time step. Ideally, the mode should be directly observed from the observations collected. Since observations are sparse in time and space, however, the mode identification is a challenging task (in addition, the knowledge of the error covariances is assumed). A heuristic can be developed to circumvent these issues, in which the gain to be used is not necessarily optimal but depends on the number and accuracy of the observations. This idea is being currently analyzed.

As mentioned before, the methods do not require the knowledge of ramp counts. It would be interesting to analyze if (and how) these counts can be derived from the density estimates obtained.

Acknowledgments

The authors would like to thank NGSIM for providing the vehicle trajectories used in this study and Saneesh Apte, Jeff Ban and Ryan Herring from CCIT for the help in processing this data as well as the Mobile Century data. The Mobile Century data collection system was built by a Nokia-UC Berkeley-Rutgers University team led by Dr. Quinn Jacobson, Professor Alexandre Bayen and Professor Marco Gruteser. The following contributors to this system are gratefully acknowledged: Ken Tracton, Toch Iwuchukwu, Dave Sutter, Baik Ho, Dan Work, Jeff Ban and Ryan Herring. Funding from Caltrans, Center for Future Urban Transport at UC Berkeley (a Volvo International Center of Excellence), Citris, Nokia, Tekes and UCTC is gratefully acknowledged.

Appendix A

This appendix provides a numerical example on the effect of the nudging factor. For illustration purposes, assume the freeway section has been divided into cells of length $\Delta x = 0.02$ miles and time into time steps of length $\Delta t = 2$ s. Let us assume that the estimated density using the model at cell i at time step h is $\hat{k}_i^h = 70$. An observation is available at the same

Table A.1
Values taken by the nudging factor for a specific case.

		Time steps affected		
		$h+1$	$h+2$	Σ
Cells affected	$i-1$	0.039	0.024	**0.063**
	$i-2$	0.054	0.033	**0.087**
	i	0.061	0.037	**0.098**
	$i+1$	0.054	0.033	**0.087**
	$i+2$	0.039	0.024	**0.063**
	Σ	**0.247**	**0.151**	**0.398**

location and at the same time, $k_i^{o,h} = 100$. Since $\hat{k}_i^h < k_i^{o,h}$, vehicles will be added, and the nudging factor will determine how many.

Let us also assume that $X_{\text{nudge}} = 0.04$ miles and $T_d = 4$ s. That is, observation from cell i at time step h affects two cells around cell i during the next two time steps (i.e. all the cells between $i-2$ and $i+2$, both inclusive, for time steps $h+1$ and $h+2$). Table A.1 shows the values taken by the nudging factor λ for the affected cells during the next two time steps (assuming $T_a = 10$ s). It can be seen how the effect of the observation from cell i at time step h reaches its maximum at $\lambda(0, \Delta t)$, and decreases with distance and time.

The total number of vehicles added because of the observation from cell i at time step h is $0.398 \cdot \left(k_i^{o,h} - \hat{k}_i^h\right) = 11.9$ vehicles. According to Eq. (12), different values for X_{nudge}, T_d, and T_a would affect in a different way neighboring cells, adding a different number of vehicles in the end. Remember that other observations are also adding (removing) vehicles to (from) the affected cells during the next two time steps.

References

Anthes, R., 1974. Data assimilation and initialization of hurricane prediction models. Journal of the Atmospheric Sciences 31 (3), 702–719.

Aw, A., Rascle, M., 2000. Resurrection of "second order" models of traffic flow. SIAM Journal on Applied Mathematics 60 (3), 916–938.

Bardos, C., Le Roux, A., Nedelec, J., 1979. First order quasilinear equations with boundary conditions. Communications in Partial Differential Equations 4 (9), 1017–1034.

Bar-Gera, H., 2007. Evaluation of a cellular phone-based system for measurements of traffic speeds and travel times: a case study from Israel. Transportation Research Part C 15 (6), 380–391.

Bar-Shalom, H., Li, X.R., 1993. Estimation and Tracking: Principles, Techniques, and Software. Artech House, Inc., Norwood, MA.

Cassidy, M., 1998. Bivariate relations in nearly stationary highway traffic. Transportation Research Part B 32 (1), 49–59.

Chu, L., Oh, S., Recker, W., 2005. Adaptive Kalman filter based freeway travel time estimation. In: Transportation Research Board 84th Annual Meeting, Washington, DC.

Chung, K., Rudjanakanoknad, J., Cassidy, M., 2007. Relation between traffic density and capacity drop at three freeway bottlenecks. Transportation Research Part B 41 (1), 82–95.

Daganzo, C., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B 28 (4), 269–287.

Daganzo, C., 1995. The cell transmission model, part II: network traffic. Transportation Research Part B 29 (2), 79–93.

Gazis, D., Knapp, C., 1971. On-line estimation of traffic densities from time-series of flow and speed data. Transportation Science 5 (3), 282–301.

Godunov, S., 1959. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. Matematicheskii Sbornik 47 (3), 271–306.

Gomes, G., Horowitz, R., 2006. Optimal freeway ramp metering using the asymmetric cell transmission model. Transportation Research Part C 14 (4), 244–262.

Hoh, B., Gruteser, M., Herring, R., Ban, J., Work, D., Herrera, J., Bayen, A., Annavaram, M., Jacobson, Q., 2008. Virtual Trip Lines for Distributed Privacy-Preserving Traffic Monitoring. MobiSys 2008, Breckenridge, CO.

Ishikawa, Y., Awaji, T., Akimoto, K., 1996. Successive correction of the mean sea surface height by the simultaneous assimilation of drifting buoy and altimetric data. Journal of Physical Oceanography 26 (11), 2381–2397.

Krause, A., Horvitz, E., Kansal, A., Zhao, F., 2008. Toward community sensing. In: ACM/IEEE International Conference on Information Processing in Sensor Netwroks (IPSN), St. Louis, MO.

Le Floch, P., 1988. Explicit formula for scalar non-linear conservation laws with boundary condition. Mathematical Methods in the Applied Sciences 10 (3), 265–287.

LeVeque, R., 2002. Finite Volume Methods for Hyperbolic Problems. Cambridge University Press, Cambridge, UK.

Lighthill, M., Whitham, G., 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. Proceedings of the Royal Society, A 229 (1178), 317–345.

Muñoz, L., Sun, X., Horowitz, R., Alvarez, L., 2003. Traffic density estimation with the cell transmitssion model. In: Proceedings of the 2003 American Control Conference, Denver, CO, pp. 3750–3755.

Nanthawichit, C., Nakatsuji, T., Suzuki, H., 2003. Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway. Transportation Research Record 1855, 49–59.

Oleinik, O.A., 1957. Discontinuous solutions of non-linear differential equations. Uspekhi Matematicheskikh Nauk 12 (3), 3–73.

Paniconi, C., Marrocu, M., Putti, M., Verbunt, M., 2003. Newtonian nudging for a Richards equation-based distributed hydrological model. Advances in Water Resources 26 (2), 161–178.

Richards, P., 1956. Shock waves on the highway. Operations Research 4 (1), 42–51.

Sanwal, K., Walrand, J., 1995. Vehicles as Probes. California PATH Working Paper UCB-ITS-PWP-95-11, Institute of Transportation Studies, University of California, Berkeley.

Schrank, D., Lomax, T., 2007. The 2007 Urban Mobility Report. Texas Transportation Institute, The Texas A&M University System, September.

Strub, I., Bayen, A., 2006. Weak formulation of the boundary conditions for scalar conservation laws: an application to highway traffic modeling. International Journal of Robust and Nonlinear Control 16 (16), 733–748.

Sun, X., Muñoz, L., Horowitz, R. 2003. Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control. In: Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, HI, pp. 6333–6338.
Sun, X., Muñoz, L., Horowitz, R., 2004. Mixture Kalman filter based highway congestion mode and vehicle density estimator and its application. In: Proceedings of the 2004 American Control Conference, Boston, MA, pp. 2098–2103.
Szeto, M., Gazis, D., 1972. Application of Kalman filtering to the surveillance and control of traffic systems. Transportation Science 6 (4), 419–439.
Treiber, M., Helbing, D., 2002. Reconstructing the spatio-temporal traffic dynamics from stationary detector data. Cooper@tive Tr@nsportation Dyn@mics 1 (3), 3.1–3.24.
Westerman, M., Litjens, R., Linnartz, J.-P., 1996. Integration of Probe Vehicle and Induction Loop Data-estimation of Travel Times and Automatic Incident Detection. PATH Research Report UCB-ITS-PRR-96-13, Institute of Transportation Studies, University of California, Berkeley.
Ygnace, J-L., Drane, C., Yim, Y., de Lacvivier, R., 2000. Travel Time Estimation on the San Francisco Bay Area Network Using Cellular Phones as Probes. California PATH Working Paper UCB-ITS-PWP-2000-18, Institute of Transportation Studies, University of California, Berkeley.
Zhang, H.M., 1998. A theory of nonequilibrium traffic flow. Transportation Research Part B 32 (7), 485–498.