



## Second-order models and traffic data from mobile sensors



Benedetto Piccoli<sup>a</sup>, Ke Han<sup>b,\*</sup>, Terry L. Friesz<sup>c</sup>, Tao Yao<sup>c</sup>, Junqing Tang<sup>b</sup>

<sup>a</sup>Department of Mathematics, Rutgers University, Camden, NJ 08102, USA

<sup>b</sup>Department of Civil and Environmental Engineering, Imperial College London, London SW7 2BU, UK

<sup>c</sup>Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, PA 16802, USA

### ARTICLE INFO

#### Article history:

Received 24 August 2013

Received in revised form 28 December 2014

Accepted 28 December 2014

#### Keywords:

Mobile sensing

Phase transition model

Higher-order traffic quantity

Emission rate

### ABSTRACT

Mobile sensing enabled by GPS or smart phones has become an increasingly important source of traffic data. For sufficient coverage of the traffic stream, it is important to maintain a reasonable penetration rate of probe vehicles. From the standpoint of capturing higher-order traffic quantities such as acceleration/deceleration, emission and fuel consumption rates, it is desirable to examine the impact on the estimation accuracy of sampling frequency on vehicle position. Of the two issues raised above, the latter is rarely studied in the literature. This paper addresses the impact of both sampling frequency and penetration rate on mobile sensing of highway traffic. To capture inhomogeneous driving conditions and deviation of traffic from the equilibrium state, we employ the second-order *phase transition model* (PTM). Several data fusion schemes that incorporate vehicle trajectory data into the PTM are proposed. And, a case study of the NGSIM dataset is presented which shows the estimation results of various Eulerian and Lagrangian traffic quantities. The findings show that while first-order traffic quantities can be accurately estimated even with a low sampling frequency, higher-order traffic quantities, such as acceleration, deviation, and emission rate, tend to be misinterpreted due to insufficiently sampled vehicle locations. We also show that a correction factor approach has the potential to reduce the sensing error arising from low sampling frequency and penetration rate, making the estimation of higher-order quantities more robust against insufficient data coverage of the highway traffic.

© 2015 Elsevier Ltd. All rights reserved.

### 1. Introduction

With the increased availability of mobile traffic data and the advancement of sensing technology, data collected through GPS, smart phones or other mobile devices have become a major source of traffic information for various applications. Advantages of mobile sensing, in comparison with fixed-location sensing (e.g. using loop detectors and cameras), include potentially complete spatial and temporal coverage of traffic network and high positioning accuracy (Herrera et al., 2010).

Traffic data related to speed, density, queue size and travel time, which are categorized as *lower-order* quantities, can often be estimated in conjunction with first-order traffic flow models such as the Lighthill–Whitham–Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956) and the cell transmission model (CTM) (Daganzo, 1994). Strub and Bayen (2006) employ a weak formulation of boundary conditions for the LWR model based on the Godunov scheme, which is then

\* Corresponding author.

E-mail addresses: [piccoli@camden.rutgers.edu](mailto:piccoli@camden.rutgers.edu) (B. Piccoli), [k.han@imperial.ac.uk](mailto:k.han@imperial.ac.uk) (K. Han), [tfriesz@psu.edu](mailto:tfriesz@psu.edu) (T.L. Friesz), [tyy1@engr.psu.edu](mailto:tyy1@engr.psu.edu) (T. Yao), [junqing.tang13@imperial.ac.uk](mailto:junqing.tang13@imperial.ac.uk) (J. Tang).

applied to the I-80 highway dataset. In [Christiani et al. \(2010\)](#), the LWR model is discretized in connection with initial and boundary conditions, which is applied to traffic estimation on a circular urban motorway using mobile data. [Claudel and Bayen \(2011\)](#) propose convex formulations for data assimilation using both Eulerian (fixed) and Lagrangian (mobile) traffic data based on the Hamilton–Jacobi representation of highway traffic and the generalized Lax–Hopf formula. [Work et al. \(2010\)](#) employ a velocity-based LWR model with transformed fundamental diagram to perform data fusion using ensemble Kalman filter. Independently, [Yuan et al. \(2011\)](#) consider the LWR model with a transformed coordinate system, namely the Lagrangian coordinates, and perform traffic estimation using extended Kalman filter. These studies mainly focus on freeway traffic.

In another line of research, mobile data have been used extensively for estimating queue size and delay at signalized intersections in arterial networks. [Ban et al. \(2009\)](#) use sampled vehicle travel times to estimate delay patterns near a signalized junction, where the authors use the LWR theory with a triangular fundamental diagram to express the relationship among flow, shock speed, queue size and queuing time. Following this work, [Ban et al. \(2011\)](#) devise a reverse modeling process that construct the dynamic queue length in real time. [Cheng et al. \(2012\)](#) further explore probe vehicle trajectories in estimating queue size in real time with the benefit of less communication cost in data collection. [Comert and Cetin \(2009\)](#) propose a statistical approach for estimating queue length using an analytical formulation based on conditional probabilities. The authors also address the estimation accuracy with a wide range of probe penetration rates. [Argote et al. \(2011\)](#) consider several *measures of effectiveness* and estimation methods to identify proper penetration rates.

Although the LWR model and the CTM have been used effectively in estimating lower-order quantities, they have been used less frequently in estimating higher-order traffic quantities such as acceleration/deceleration, deviation (perturbation), emission and fuel consumption rates. There exist a number of attempts to estimate acceleration/deceleration or emission rates through differentiating macroscopic traffic quantities analytically or numerically (e.g. [Luspay et al., 2010](#)). However, in this process higher-order variations inherent in these quantities, typically on a microscopic scale, are insufficiently captured due to the low temporal-spatial resolution of the traffic data and the discrete models. To fill this gap, this paper proposes a second-order traffic flow model supported by high-resolution mobile data to address the issue of estimating high-order traffic quantities. Unlike most existing studies on mobile sensing which primarily focus on probe penetration rate ([Demers et al., 2006](#); [Kwon et al., 2007](#); [Yim and Cayford, 2001](#)), we consider the additional effect of under sampling on the estimation accuracy. This is a concern because most mobile data provide location or speed of a moving vehicle every 3–4 s (such as GPS data), but higher-order variations in speed, acceleration and emission may take place on a smaller time scale; this is true especially for congested and unstable traffic. Such an observation has urged the need to examine the efficacy of existing sensing technologies and estimation methods in reconstructing the profiles for these higher-order quantities.

The second-order traffic flow model employed by this paper is the hyperbolic *phase transition model* (PTM), first introduced by [Colombo \(2002\)](#) and studied subsequently by [Colombo and Corli \(2002\)](#), [Colombo et al. \(2007\)](#) and [Blandin et al. \(2012\)](#). Second- and higher-order traffic models were proposed by many researchers to overcome some limitations of the LWR model in describing complex waves observed in vehicular traffic. Most second-order models tend to pose, in addition to the LWR-type equation for the conservation of vehicles, a second equation for the conservation or balance of momentum. One of the first such models is by Payne and Whitham in the 1970s ([Payne, 1971](#); [Whitham, 1974](#)). In a celebrated paper ([Daganzo, 1995](#)) Daganzo criticized second-order models by showing various drawbacks including the possibility of cars going backward. Most of such drawbacks were later addressed by the Aw–Rascle–Zhang model, independently proposed by [Aw and Rascle \(2000\)](#) and [Zhang \(2002\)](#). More recently, the phase transition models drew increased attention from researchers for their capability of representing complex waves while keeping the LWR structure for light traffic, see [Blandin et al. \(2011\)](#) for more detailed discussion.

Let us further comment on the possible use of second-order model for urban arterial traffic. The aforementioned complex wave phenomenon, well captured by second-order models, are mainly observed in highway traffic. Indeed, phantom waves, stop-and-go waves and others need a long stretch of road with no interruption to manifest themselves. The situation of arterial traffic is quite different because of the presence of many junctions with traffic signs or signals. Since the LWR model captures well backward wave propagation from junctions or signals, and vehicle movements on arterials are relatively uniform, it is typically sufficient to describe arterial traffic, while very limited to model more complex waves in highway traffic.

In order to address issues related to higher-order traffic quantities, we focus on the *Next Generation SIMulation* (NGSIM) dataset.<sup>1</sup> The NGSIM program collected high-quality traffic and vehicle trajectory data on a stretch of I-80 highway in California. A total of 45 min of traffic data were collected, segmented into three 15 min periods. The dataset contains vehicle trajectory recorded at a high resolution of every 0.1 s. Derived information on instantaneous velocity and acceleration is also available. A detailed description of the NGSIM field experiment and data will be provided in Section 5.

### 1.1. Contribution and findings

This paper addresses the effectiveness of mobile sensing and traffic estimation taking into account two main factors: sampling frequency and probe penetration rate. The subject of estimation will include both lower- and higher-order traffic quantities. To support this study, we employ the second-order phase transition model (PTM) as well as its modifications. The

<sup>1</sup> <http://ngsim-community.org/>.

underpinning dataset is provided by the NGSIM program, which contains high resolution vehicle trajectories on a segment of I-80. Such detailed vehicle trajectories provide unique information regarding the higher-order variations of traffic, which is unavailable through traditional mobile or fixed sensors. We propose several data fusion schemes that integrate mobile data into the PTM under various assumptions. As a result, vehicle speed, acceleration, density, deviation (perturbation), engine power demand and emission rate can be estimated along vehicle trajectories (these quantities are Lagrangian as they are associated with moving vehicles). We demonstrate how these estimations deteriorate with less frequent sampling of vehicle locations, with special attention given to the difference between lower- and higher-order quantities. We also present a method for estimating Eulerian traffic quantities; i.e. quantities associated with a give point in the temporal-spatial domain. The joint effects of sampling frequency and probe penetration rate in reconstructing lower- and higher-order quantities are assessed. It is found that, for both Lagrangian and Eulerian estimations, lower-order traffic variables (such as density) can be estimated relatively accurately, while higher-order ones (such as emission and perturbation) are far more susceptible to under sampling and lower penetration rate. This is consistent with the expectation that higher-order variations tend to be overlooked by data presented at a low resolution. We also show that Eulerian estimation of higher-order quantities enjoy improved accuracy than Lagrangian estimation due to the effect of spatial-temporal aggregation. And, the efficacy of Eulerian estimation can be further improved by applying a correction factor approach. This is demonstrated by using hydrocarbon emission rate as an example and proven to be an effective estimation method even with low sampling frequency and penetration rate.

The main contributions and findings made by this paper are summarized as follows:

- Several methods for integrating vehicle trajectory data into the phase transition model are proposed. These methods are further extended to perform both Lagrangian and Eulerian estimation of traffic quantities.
- Along each vehicle trajectory, the estimation of higher-order traffic quantities deteriorates significantly when the sampling frequency decreases, while the estimation of first-order quantities remain relatively accurate with the same sampling frequency.
- For the Eulerian estimation, we provide numerical results on density (first-order) and HC emission rate (higher-order) with varying sampling frequency and penetration rate. The HC estimation enjoys an improved accuracy compared to the Lagrangian estimation due to the averaging of multiple measurements, although it also deteriorates with under sampling and lower penetration rate.
- We provide a comparative study of the LWR model (first order) and the phase transition model in terms of estimating first-order density (notice that higher-order quantities such as acceleration and perturbation are not captured explicitly by the LWR model). The PTM outperforms the LWR model in terms of estimation accuracy. In addition, deviation of individual vehicles from the equilibrium state, which is not captured by the LWR model, has been illustrated numerically.
- A correction factor approach is proposed for estimating HC emission rate with insufficient mobile data coverage. We employ a regression model to correct the predicted emission rate in order to minimize the misinterpretation of the ground truth caused by lower sampling frequencies or probe penetration rates. The validity and effectiveness of this approach is shown through cross validation.

The rest of this paper is organized as follows: Section 2 recaps several numerical approximations of basic traffic quantities based on vehicle trajectory data. Section 3 introduces the phase transition model (PTM), followed by three data fusion schemes. Models for HC emission rate and engine power demand are also described. Section 4 illustrates the procedure employed to reduce random noises in the raw dataset and errors derived from numerical differentiation. In Section 5, we calibrate the PTM using the NGSIM data. Section 6 assesses the estimation quality for first- and second-order quantities along vehicle trajectories, when the sampling frequency varies. Sections 7 and 8 perform Eulerian estimation for density and HC emission rate respectively, and evaluate its effectiveness against various sampling frequencies and penetration rates. A correction factor approach for improving the accuracy and reliability of HC emission is also presented and tested. Section 9 provides some concluding remarks.

## 2. Estimating basic traffic quantities

Onboard sensors such as GPS and smart phone measure the location of a moving car every  $\delta t$  seconds, where  $\delta t$  is related to the device's characteristics such as desired precision and transmission capacity. For a given vehicle, we denote by  $x(t)$  its location and by  $v(t)$  its velocity at time  $t$ . Assume that the location is recorded at three consecutive time stamps  $t_1, t_2$  and  $t_3$  with  $t_2 - t_1 = t_3 - t_2 = \delta t$ . From these measurements one can deduce approximate velocities in the time intervals  $[t_1, t_2]$  and  $[t_2, t_3]$  respectively as

$$v_{1,2} = \frac{x(t_2) - x(t_1)}{\delta t}, \quad v_{2,3} = \frac{x(t_3) - x(t_2)}{\delta t} \quad (2.1)$$

The velocity at time  $t_2$  is approximated as

$$v(t_2) \approx \frac{v_{1,2} + v_{2,3}}{2} = \frac{x(t_3) - x(t_1)}{2\delta t} \quad (2.2)$$

One also gets estimate for the acceleration:

$$a(t_2) = \frac{D}{Dt} v(t, x) \Big|_{t=t_2} \approx \frac{v_{2,3} - v_{1,2}}{\delta t} = \frac{x(t_3) - 2x(t_2) + x(t_1)}{\delta t^2} \quad (2.3)$$

where  $v(t, x)$  denotes the velocity function in the Eulerian coordinates;  $D/Dt = d/dt + v \cdot d/dx$  represents the material derivative in the Eulerian coordinates corresponding to the acceleration of the car in the Lagrangian coordinate.

Another important quantity to estimate is the spatial variation of velocity in the Eulerian coordinate. For notation convenience, we set  $x_i \doteq x(t_i)$ . Assuming a mild variation in time of the Eulerian velocity  $v(t, x)$ , we write:

$$v\left(t_2, \frac{x_2 + x_1}{2}\right) \approx v_{1,2}, \quad v\left(t_2, \frac{x_3 + x_2}{2}\right) \approx v_{2,3},$$

from which we get, by setting  $\delta x = \frac{x_3 + x_2}{2} - \frac{x_2 + x_1}{2}$ , that

$$\frac{\partial}{\partial x} v\left(t_2, \frac{x_3 + x_2}{2} + \frac{x_2 + x_1}{2}\right) \approx \frac{v_{2,3} - v_{1,2}}{\delta x}$$

In other words,

$$\frac{\partial}{\partial x} v\left(t_2, \frac{x_3 + 2x_2 + x_1}{4}\right) \approx \frac{v_{2,3} - v_{1,2}}{\frac{x_3 - x_1}{2}} = \frac{2}{\delta t} \frac{x_3 - 2x_2 + x_1}{x_3 - x_1} \quad (2.4)$$

Clearly such approximation is acceptable as long as the variation between  $v_{1,2}$  and  $v_{2,3}$  is not too large.

### 3. Model fitting using vehicle trajectory data

We consider the phase transition model (PTM) as well as some of its variations to represent the dynamics of highway traffic. In addition, we also consider models proposed by Barth et al. (2007) and Ahn et al. (2002) to describe instantaneous engine power demand and emission rate based on detailed vehicle trajectory. All these models will be described in detail in this section.

#### 3.1. The phase transition model

The hyperbolic phase transition model (PTM) for traffic flow belongs to a class known as *second-order*, since it captures second-order variations of traffic in addition to the average velocity and density. Other second-order models include the Payne–Whitham model proposed independently by Payne (1971, 1979) and Whitham (1974), and the Aw–Rascle–Zhang model developed by Aw and Rascle (2000) and Zhang (2002). The phase transition model is motivated by the empirical observation that when the vehicle density exceeds certain critical value, the density–flow pairs are scattered in a two-dimensional region instead of forming a one-to-one relationship. This is in contrast to the first-order kinematic wave models such as the classical Lighthill–Whitham–Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956).

The phase transition model consists of two distinct traffic phases: the uncongested phase and the congested phase. In the uncongested phase, the dynamic is governed by the LWR model

$$\begin{cases} \partial_t \rho(t, x) + \partial_x [\rho(t, x) \cdot v(t, x)] = 0 \\ v(t, x) = v(\rho(t, x)) \end{cases} \quad (3.5)$$

where  $\rho(t, x)$  denotes density, and the velocity  $v(\rho(t, x))$  is expressed as a function of density. The product  $\rho(t, x) \cdot v(t, x)$  represents flow. On the other hand, the congested phase is governed by the following system of conservation laws:

$$\begin{cases} \partial_t \rho(t, x) + \partial_x [\rho(t, x) \cdot v(t, x)] = 0 \\ \partial_t q(t, x) + \partial_x [(q(t, x) - q^*) \cdot v(t, x)] = 0 \\ v(t, x) = v(\rho(t, x), q(t, x)) \end{cases} \quad (3.6)$$

where the velocity  $v(\rho, q)$  depends not only on the local density  $\rho$ , but also on  $q$  which describes the perturbation or deviation from the equilibrium state.  $q^*$  is a given constant corresponding to the equilibrium (stationary) traffic state. Some choices of the functional form  $v(\rho, q)$  include

$$v(\rho, q) = \left(1 - \frac{\rho}{\rho_{jam}}\right) \cdot \frac{q}{\rho} \quad \text{or} \quad v(\rho, q) = v(\rho)(1 + q) \quad (3.7)$$

where  $\rho_{jam}$  denotes the jam density,  $v(\rho)$  denotes the equilibrium velocity function. In this paper, we employ the following functional form for the velocity.

$$v(\rho, q) = A(\rho_{jam} - \rho) + B(q - q^*)(\rho_{jam} - \rho) \quad (3.8)$$

for some positive parameters  $A, B$  and  $\rho_{jam}$ , where  $q - q^* \in [-1, 1]$ .<sup>2</sup>

**Remark 3.1.** Several factors should be taken into account when formulating an appropriate functional form for  $v(\rho, q)$ :

- (i) When  $q = q^*$  (the equilibrium state), the velocity  $v(\rho, q)$  coincides with the stationary density–velocity relationship.
- (ii)  $v(\rho, q) \geq 0$ , and zero velocity arises only at densities corresponding to the jam density in the equilibrium state (this property is posed by Garavello and Piccoli (2009) to ensure physically correct behavior of models on networks).
- (iii) Assume that  $q - q^* \in [q_{min}, q_{max}]$ . The 2-D region formed by the curves  $v(\rho, q^* + q_{min})$  and  $v(\rho, q^* + q_{max})$  should match the empirical density–velocity data plot (see Fig. 5 for an example).

As we show later in the model calibration (Section 5), the density–velocity plot suggests a well-defined affine stationary relationship for the congested branch, which is represented by the first term on the right hand side of (3.8). In addition, the upper and lower envelopes of the 2-D region, expressed respectively as  $(A + B)(\rho_{jam} - \rho)$  and  $(A - B)(\rho_{jam} - \rho)$ , also match well with the data points; see Fig. 5. Finally, the proposed function (3.8) clearly satisfies requirements (i) and (ii). This justifies our choice of  $v(\rho, q)$ .

For the PTM in general, there is some flexibility in choosing the system of equations for the congested phase. For instance, Goatin (2006) proposes a phase transition model which employs the Aw–Rascle–Zhang equations (Aw and Rascle, 2000; Zhang, 2002) for the congested phase. Furthermore, one may consider a version of the PTM by taking into account the reaction time of drivers. More specifically, following Siebel and Mauser (2006), we write

$$\partial_t q(t, x) + \partial_x [(q(t, x) - q^*) \cdot v(t, x)] = \frac{q^* - q(t, x)}{T - \tau} \quad (3.9)$$

as the second equation for the congested phase. The non-vanishing right hand side is called the Siebel–Mauser source term. Here,  $\tau$  (in second) is a reaction time and typically varies within  $[0.5, 2]$  (Koppa, 1997). In this paper, the value of  $\tau$  is set to be 1 (second). Following Siebel and Mauser (2006) and Kühne and Michalopoulos (1997), we choose  $T = 2/3$  (second) and therefore  $T - \tau = -1/3$ . Finally  $T - \tau$  is modeled as a factor which can be positive (for very small or very high densities) or negative (for intermediate densities). One should realize that a negative factor gives rise to stable traffic, while a positive one produces instabilities (Siebel and Mauser, 2006).

### 3.2. Estimating traffic quantities using the PTM

The primary variables of the PTM are vehicle density  $\rho$  (first-order) and deviation  $q$  (second-order). It is possible to estimate both quantities within the PTM using just vehicle trajectory data, as we demonstrate below. Note that subsequent derivations are established based on the congested phase of the PTM for two reasons: (1) the NGSIM dataset suggests a prevailing congested condition of traffic during the time of study; (2) it is mainly within the congested phase where traffic instabilities and heterogeneous driving conditions may arise.

We propose three methods to perform the estimation of the two variables in the PTM. These methods rely on different assumptions made on the model and the traffic conditions.

#### 3.2.1. Phase transition model with source term: Method 1

We consider the PTM with velocity in the congested phase expressed as

$$v(\rho, q) = A(\rho_{jam} - \rho) + B(q - q^*)(\rho_{jam} - \rho) \quad (3.10)$$

In addition, we add a Siebel–Mauser type source term. Thus the equation for the congested phase becomes

$$q_t + [(q - q^*) \cdot v]_x = \frac{q^* - q}{T - \tau} \quad (3.11)$$

where, without causing any confusion, we simplify the notations by dropping  $(t, x)$  and using subscripts  $t$  and  $x$  to denote partial derivatives with respect to time and location. Such notation convention will be adopted throughout the rest of this paper.

Method 1 employs the following simplifications.

**Assumption 1.** Assume small variations of  $\rho$  in  $t$ , and  $\rho, q$  in  $x$ . That is,  $\rho_t, \rho_x, q_x \approx 0$  with only  $q_t$  being non-negligible.

By Assumption 1, (3.11) reduces to

$$q_t = \frac{q^* - q}{T - \tau} \quad (3.12)$$

Then we can derive

<sup>2</sup> In general, the deviation  $q - q^* \in [q_{min}, q_{max}]$  for some lower and upper bounds  $q_{min} < 0$  and  $q_{max} > 0$ . This paper assumes that the deviations are symmetric around the equilibrium state; that is  $q - q^* \in [-q_{max}, q_{max}]$  for some  $q_{max} > 0$ . Without loss of generality, we further assume that  $q - q^* \in [-1, 1]$  since  $q_{max}$  can be factored into the coefficient  $B$ , which will later be calibrated using field data in Section 5.

$$\begin{aligned} \frac{D}{Dt} v &= \partial_t v(\rho(t, x), q(t, x)) + v \cdot \partial_x v(\rho(t, x), q(t, x)) = v_\rho \cdot \rho_t + v_q \cdot q_t + v \cdot (v_\rho \cdot \rho_x + v_q \cdot q_x) \approx v_q \cdot q_t \\ &= B(\rho_{jam} - \rho) \cdot \frac{q^* - q}{T - \tau} = \frac{1}{T - \tau} B(\rho_{jam} - \rho)(q^* - q) = \frac{1}{T - \tau} (A(\rho_{jam} - \rho) - v) \end{aligned} \quad (3.13)$$

Taking into account only the measurements of  $v$  and  $Dv/Dt$ , and by introducing variables  $\hat{\rho} \doteq \rho_{jam} - \rho$ ,  $\hat{q} \doteq q - q^*$ , we deduce from (3.13) that

$$\hat{\rho} = (\rho_{jam} - \rho) = \frac{1}{A} \left( v + (T - \tau) \frac{Dv}{Dt} \right) \quad (3.14)$$

$$\hat{q} = (q - q^*) = -\frac{A(T - \tau)}{B} \frac{\frac{Dv}{Dt}}{v + (T - \tau) \frac{Dv}{Dt}} \quad (3.15)$$

Following our discussion at the end of Section 3.1, we take  $T - \tau = -\frac{1}{3}$ . The velocity  $v$  is estimated according to (2.2), and the acceleration  $Dv/Dt$  is computed from (2.3).

### 3.2.2. Phase transition model with source term: Method 2

The second method again assumes a Siebel–Mauser source term, and relies on the following assumption.

**Assumption 2.** Assume that  $\rho_t, q_x \approx 0$ .

Then Eq. (3.11) becomes

$$q_t + v_x(q - q^*) = \frac{q^* - q}{T - \tau} \quad (3.16)$$

We can now write

$$\frac{Dv}{Dt} = v_t + v v_x \approx v_q q_t + v v_x = v_q \left( \frac{q^* - q}{T - \tau} - (q - q^*) v_x \right) + v v_x \quad (3.17)$$

Recalling the variables  $\hat{\rho} \doteq \rho_{jam} - \rho$ ,  $\hat{q} \doteq q - q^*$ , we obtain

$$v = A(\rho_{jam} - \rho) + B(q - q^*)(\rho_{jam} - \rho) = \hat{\rho}(A + B\hat{q}) \quad (3.18)$$

and deduce from (3.17) and (3.18) that

$$\frac{Dv}{Dt} - v v_x = -v_q \left( \frac{\hat{q}}{T - \tau} + \hat{q} v_x \right) = -B\hat{\rho} \left( \frac{\hat{q}}{T - \tau} + \hat{q} v_x \right) \quad (3.19)$$

From (3.18) and (3.19) we immediately get the expressions for  $\hat{\rho}$  and  $\hat{q}$  in terms of  $v$ ,  $v_x$  and  $\frac{Dv}{Dt}$ .

$$\hat{\rho} = \frac{1}{A} \frac{v + (T - \tau) \frac{Dv}{Dt}}{1 + (T - \tau) v_x} \quad (3.20)$$

$$\hat{q} = \frac{A}{B} \frac{(T - \tau)(v v_x - \frac{Dv}{Dt})}{v + (T - \tau) \frac{Dv}{Dt}} \quad (3.21)$$

The quantities  $\frac{Dv}{Dt}$ ,  $v_x$  are given by (2.3) and (2.4) respectively. Regarding velocity  $v$ , notice that if one approximates it with (2.1), then

$$\frac{Dv}{Dt} - v v_x \approx \frac{x(t_{i+1}) - 2x(t_i) + x(t_{i-1}))}{\delta t^2} - \frac{x(t_{i+1}) - x(t_{i-1}))}{2\delta t} \cdot \frac{2}{\delta t} \cdot \frac{x(t_{i+1}) - 2x(t_i) + x(t_{i-1}))}{x(t_{i+1}) - x(t_{i-1}))} \equiv 0 \quad (3.22)$$

which renders (3.21) identically zero. To avoid such a trivial case, one should instead approximate  $v$  by

$$v(t_2) \approx v_{1,2} = \frac{x(t_2) - x(t_1)}{\delta t}, \quad \text{or} \quad v(t_2) \approx v_{2,3} = \frac{x(t_3) - x(t_2)}{\delta t}, \quad (3.23)$$

or use robust numerical differentiation illustrated later in Section 4.2.

### 3.2.3. Phase transition model without source term: Method 3

We consider the original PTM without the Siebel–Mauser type source term:

$$\begin{cases} \rho_t + (\rho \cdot v)_x = 0 \\ q_t + ((q - q^*) \cdot v)_x = 0 \\ v(\rho, q) = A(\rho_{jam} - \rho) + B(q - q^*)(\rho_{jam} - \rho) \end{cases} \quad (3.24)$$

Computation in this case is less straightforward than the previous two as no simplification is made in this case. We start with the identity

$$\frac{Dv}{Dt} = v_t + v \cdot v_x = v_\rho \rho_t + v_q q_t + v(v_\rho \rho_x + v_q q_x)$$

Using (3.24), we have

$$\begin{aligned} \frac{Dv}{Dt} &= v_\rho(-\rho_x v - \rho v_x) + v_q(-q_x v - (q - q^*)v_x) + v(v_\rho \rho_x + v_q q_x) = -v_x(v_\rho \rho + v_q(q - q^*)) \\ &= v_x(A\rho + B(q - q^*)(2\rho - \rho_{jam})) \end{aligned} \quad (3.25)$$

Recall the variables  $\hat{\rho} = \rho_{jam} - \rho$  and  $\hat{q} = q - q^*$ . Combining (3.25) with the expression of  $v(\rho, q)$  and solving for  $\rho$  leads to

$$\rho = \frac{1}{B\hat{q}} \left( v + \frac{1}{v_x} \frac{Dv}{Dt} - A\rho_{jam} \right) \quad (3.26)$$

One immediate observation from (3.26) is that  $\hat{q}$  and  $v + \frac{1}{v_x} \frac{Dv}{Dt} - A\rho_{jam}$  always have the same sign. Substituting (3.26) into (3.25), we get

$$B^2 \rho_{jam} \hat{q}^2 + B \left( 2A\rho_{jam} - \frac{1}{v_x} \frac{Dv}{Dt} - 2v \right) \hat{q} - A \left( v + \frac{1}{v_x} \frac{Dv}{Dt} - A\rho_{jam} \right) = 0 \quad (3.27)$$

which is a quadratic equation in the variable  $\hat{q}$ . The discriminant of such a quadratic equation is

$$\Delta = 4B^2 \left( v + \frac{1}{2} \frac{1}{v_x} \frac{Dv}{Dt} - A\rho_{jam} \right)^2 + 4AB^2 \rho_{jam} \left( v + \frac{1}{v_x} \frac{Dv}{Dt} - A\rho_{jam} \right) = 4B^2 \left( \left( v + \frac{1}{2} \frac{1}{v_x} \frac{Dv}{Dt} \right)^2 - A\rho_{jam} v \right) \quad (3.28)$$

In order for any meaningful real root of (3.27) to exist, a necessary condition is that  $\Delta$  is nonnegative; that is,

$$\left| v + \frac{1}{2} \frac{1}{v_x} \frac{Dv}{Dt} \right| \geq \sqrt{A\rho_{jam} v} \quad (3.29)$$

If the equality holds in (3.29), a real solution of  $\hat{q}$  exists if and only if

$$v + \frac{1}{v_x} \frac{Dv}{Dt} - A\rho_{jam} < 0 \quad (3.30)$$

In the case where the strict inequality in (3.29) holds, two distinct roots  $\hat{q}_1$  and  $\hat{q}_2$  exist. We distinguish between two further cases:

- If

$$v + \frac{1}{v_x} \frac{Dv}{Dt} - A\rho_{jam} > 0 \quad (3.31)$$

Then  $\hat{q}_1 \hat{q}_2 < 0$ . Eq. (3.27) has one positive root and one negative root. By (3.26), one should choose the positive root since  $\rho$  must be non-negative.

- If

$$v + \frac{1}{v_x} \frac{Dv}{Dt} - A\rho_{jam} \leq 0 \quad (3.32)$$

Then  $\hat{q}_1 \hat{q}_2 \geq 0$ . Eq. (3.27) has two roots with the same sign.<sup>3</sup> In view of (3.26) and (3.32) to ensure that  $\rho$  is nonnegative, both  $\hat{q}_1$  and  $\hat{q}_2$  should be nonpositive and at least one root is negative. This in turn requires that

$$2A\rho_{jam} - \frac{1}{v_x} \frac{Dv}{Dt} - 2v > 0 \quad (3.33)$$

In view of (3.32), a sufficient condition for (3.33) to hold is  $\frac{1}{v_x} \frac{Dv}{Dt} > 0$ .

It turns out that the above analysis can be more easily interpreted in a discrete-time and numerical setting. Let us consider three consecutive time stamps  $t_{i-1}$ ,  $t_i$ ,  $t_{i+1}$ . First, notice that

$$\frac{1}{v_x} \frac{Dv}{Dt} \approx \frac{\delta t}{2} \cdot \frac{x(t_{i+1}) - x(t_{i-1})}{x(t_{i+1}) - 2x(t_i) + x(t_{i-1}))} \cdot \frac{x(t_{i+1}) - 2x(t_i) + x(t_{i-1}))}{\delta t^2} = \frac{x(t_{i+1}) - x(t_{i-1}))}{2\delta t} \approx v(t_i) \geq 0 \quad (3.34)$$

In light of this calculation, the feasibility condition (3.29) becomes

$$\frac{3}{2} v \geq \sqrt{A\rho_{jam} v}, \quad \text{or} \quad v \geq \frac{4}{9} A\rho_{jam} \quad (3.35)$$

Using (3.34), the decision rule following (3.29) can be explicitly summarized as follows.

<sup>3</sup> Zero is considered to have both positive and negative signs.



**Algorithm 1.** PTM without source term

---

```

if  $v < \frac{4}{9}A\rho_{jam}$  then the model has no solution
else if  $v = \frac{4}{9}A\rho_{jam}$  then  $\rho = \frac{\rho_j}{3}$ ,  $\hat{q} = -\frac{A}{3B}$ 
else
  if  $v > \frac{1}{2}A\rho_{jam}$  then

    
$$\rho = \frac{2\rho_{jam}(2v - A\rho_{jam})}{3v - 2A\rho_{jam} + \sqrt{9v^2 - 4A\rho_{jam}v}}, \quad \hat{q} = \frac{3v - 2A\rho_{jam} + \sqrt{9v^2 - 4A\rho_{jam}v}}{2B\rho_{jam}}$$


  else

    
$$\rho = \frac{2\rho_{jam}(2v - A\rho_{jam})}{3v - 2A\rho_{jam} - \sqrt{9v^2 - 4A\rho_{jam}v}}, \quad \hat{q} = \frac{3v - 2A\rho_{jam} - \sqrt{9v^2 - 4A\rho_{jam}v}}{2B\rho_{jam}}$$


  end if
end if

```

---

**Remark 3.2.** We note that the above computational procedure does not produce any result if the velocity is below  $\frac{4}{9}A\rho_{jam}$ . Such an assumption may be rather restrictive in application, especially for traffic where low speed is prevailing, although this method is exact and consistent with the original phase transition model. As we later show in Table 2, Method 3 is more resilient to under sampling in estimating  $\hat{q}$  than the other two methods, whenever applicable.

### 3.3. Estimating power demand and vehicle emission rates

Car emission and/or fuel consumption rates are considered in this section. They belong to high-order quantities as they are usually expressed as nonlinear functions of instantaneous speed and acceleration (Post et al., 1984; Barth et al., 2007). That is, the emission and/or fuel consumption rates are related to the modal operation of a moving vehicle such as idle, steady-state cruise, acceleration or deceleration. It requires detailed vehicle trajectory data in order to capture these driving conditions.

It has been shown that fuel consumption is closely related to the engine power expressed as the *power demand function* (PDF) (Post et al., 1984; Barth et al., 2007), which in turn can be expressed as a nonlinear function of vehicle speed and acceleration, along with parameters related to road grade, air drag and other physical conditions. One example of such a PDF is provided in Barth et al. (2007) as

$$Z = \frac{M}{1000} \cdot v \cdot (a + g \sin \theta) + \left( MgC_r + \frac{\sigma}{2} v^2 A_c C_d \right) \cdot \frac{v}{1000} \quad (3.36)$$

where  $Z$  (in kw) is the power demand;  $v$  (in km/h) and  $a$  (in km/h/s) denote instantaneous speed and acceleration, respectively.  $M$  (in kg) is vehicle mass and is chosen uniformly as 1200 in this paper for simplicity;  $g = 9.81(\text{m/s}^2)$  is the gravitational constant;  $\theta$  is the road grade and is chosen to be zero;  $\sigma = 1.225 (\text{kg/m}^3)$  is the mass density of air. Finally,  $C_r = 0.005$ ,  $A_c = 2.6 (\text{m}^2)$ , and  $C_d = 0.3$  denote respectively the rolling resistance coefficient, the cross-sectional area, and the aerodynamic drag coefficient, and their values are suggested by Chan (1997).

Regarding vehicle emission, we consider hydrocarbon (HC) and employ the model proposed in Ahn et al. (2002). That paper presents a hybrid 3rd-order polynomial regression model with logarithmic transformation where the acceleration case ( $a \geq 0$ ) and the deceleration case ( $a < 0$ ) are fitted separately. The HC emission rate is modeled as

$$\ln r^{HC} = \begin{cases} \sum_{i=0}^3 \sum_{j=0}^3 U_{ij} \cdot v^i \cdot a^j & \text{if } a \geq 0 \\ \sum_{i=0}^3 \sum_{j=0}^3 V_{ij} \cdot v^i \cdot a^j & \text{if } a < 0 \end{cases} \quad (3.37)$$

where the matrices  $\{U_{ij}\}_{i,j=0}^3$  and  $\{V_{ij}\}_{i,j=0}^3$  are explicitly given in Ahn et al. (2002), and are omitted here. This model is shown to better approximate the HC emission for combinations of relatively high speed and acceleration, and are thus chosen here since part of the monitored traffic is highly unstable containing accelerations up to 10 km/h per second.



**Remark 3.3.** Here we assume, as in almost all macroscopic traffic models, that all vehicles have the same characteristics, e.g. weight, engine type, size, and so on. The NGSIM dataset contains vehicle type which could potentially improve the estimation accuracy. This is, however, not pursued by this paper as we aim to describe the evolution of traffic at an aggregate level with bounded granularity on the characteristics of the flows.

We further note the fact that other types of models for instantaneous emission rate, power demand function and/or fuel consumption rate can be equally applied to our framework, as long as they are expressible as functions of speed, acceleration, and vehicle-specific characteristics. Although employing these alternative models will yield quantitatively different results, the general conclusion drawn from this paper will remain valid; that is, estimations with under sampling of the vehicle location and lower probe penetration rate tend to misinterpret these quantities (see Sections 6.3 and 8.1), and the degree of such miscalculation depends on a number of factors, e.g. traffic condition (stable/unstable), vehicle type, and model parameters.

#### 4. Reconstructing ground-truth data

The NGSIM contains raw data in video format on vehicle trajectories with a high time resolution. Derived vehicle-specific quantities such as velocity and acceleration are also included therein. However, sensing errors are expected to exist, as are derived errors arising from numerical differentiation. There exist a number of ways to pre-process those vehicle trajectory data via data smoothing techniques and robust numerical differentiation; see Montanino and Punzo (2013) and Punzo et al. (2011) for examples. Based on these prior approaches, we present here two methods for reducing noises inherent in the dataset and reconstructing the ground-truth data for later use.

##### 4.1. Robust smoothing methods

Several standard smoothing techniques are available for improving regularity of the quantities of interest; e.g. the *moving average method*; the LOESS (local regression using weighted linear least squares and a second-degree polynomial model); and the LOWESS (local regression using weighted linear least square and first-degree polynomial model). We refer the reader to Cleveland et al. (1990) and Cleveland and Devlin (1988) for more elaborated discussions. However, these smoothing techniques, which are mostly based on least-square regression, are known to be sensitive to outliers, which should ideally be removed before smoothing is applied. A few generic methods for detecting outliers in a dataset are summarized by Blatná (2005). For the time series of vehicle location, velocity and acceleration in the NGSIM data, removal of the outliers is done in Montanino and Punzo (2013) and Punzo et al. (2011).

For the application described in this paper, we utilize the robust versions of the “LOESS” or “LOWESS” methods, namely “RLOESS” or “RLOWESS”, which are implemented and documented in the MATLAB Curve Fitting toolbox.<sup>4</sup> Taking “RLOESS” as an example, it first smoothes the value at a given point through local regression using nearby data points (the range of those points is determined by a parameter called *span*); this is precisely the procedure employed by “LOESS”. Additionally, the “RLOESS” calculates the residuals from the aforementioned regression and assigns a robust weight to each data point within the span:

$$w_i = \begin{cases} \left(1 - \left(\frac{r_i}{6MAD}\right)^2\right)^2 & \text{if } |r_i| < 6MAD \\ 0 & \text{if } |r_i| \geq 6MAD \end{cases} \quad (4.38)$$

where  $r_i$  denotes the residual of the  $i$ th data point, and MAD (*median absolute deviation*) is defined as  $\text{median}(|r_i|)$  over all data points. The MAD is a measure of how spread out the residuals are. If  $r_i$  for a given  $i$  is larger than  $6MAD$ , then the  $i$ th point is considered an outlier and is assigned weight 0. Finally, the regression is applied again to smooth the data, but with the robust weights (4.38).

**Remark 4.1.** If one applies the robust local regression method to location, speed, and acceleration data separately, different degree of smoothing (*span*) should be used. The reason is that the regularities of these quantities are decreasing, and with each differentiation, the error becomes magnified into increasingly larger ones. Moreover, applying the smoothing separately means that these quantities may no longer be consistent as the derivative of the smoothed location is no longer the smoothed speed, and etc. Such an inconsistency is illustrated further in Fig. 1.

##### 4.2. Regularization of numerical differentiation

It is conventionally acknowledged that finite-difference approximations, such as those expressed by (2.1) and (2.2), tend to amplify noises existing in the subject of numerical differentiation. In some applications, smooth-then-differentiate or differentiate-then-smooth may not yield satisfactory results (Chartrand, 2011). Instead of applying the smoothing method to individual trajectories of location, velocity, or acceleration, one alternative is to utilize a numerical differentiation scheme that is robust against noises in the data. One such robust scheme is Tikhonov regularization (Tikhonov, 1963; Chartrand,

<sup>4</sup> URL: <http://www.mathworks.com/help/curvefit/smoothing-data.html>.

2011), and it works as follows. Given a function  $f$ , its derivative on  $[0, T]$ , denoted by  $g$ , is the solution of the following minimization problem.

$$\min_g \mathcal{F}(g) = \alpha R(g) + DF(\mathcal{I}g - f) \quad (4.39)$$

where  $R(g)$  is a regularization term that penalizes the irregularity in  $g$ , and it is chosen as the square of the  $L^2$  norm of the derivative of  $g$  in our study:

$$R(g) \doteq \int_0^T |g'(\tau)|^2 d\tau$$

Such a choice forces  $g$  to be continuous and reasonably regular by penalizing large values of  $|g'|$ . The operator  $\mathcal{I}$  represents integration, and  $\mathcal{I}g(t) \doteq \int_0^t g(\tau) d\tau$ ,  $t \in [0, T]$ .  $DF(\mathcal{I}g - f)$  is a *data fidelity* term that penalizes the discrepancy between  $\mathcal{I}g$  and  $f$ , and it is chosen as the square of the  $L^2$  norm of  $\mathcal{I}g - f$ .  $\alpha > 0$  is a weighting parameter that balances between the two terms.

In a discrete-time setting, denote by  $f_i$  and  $g_i$  the discrete values of  $f$  and  $g$  respectively,  $i = 1, \dots, n$ . The integration operator is approximated by a rectangular or trapezoidal quadrature and can be written as an affine transformation<sup>5</sup> in discrete time:

$$\{(\mathcal{I}g)_i : 1 \leq i \leq n\} = Ag + b \quad A \in \mathbb{R}^{n \times n}, \quad b \in \mathbb{R}^n$$

where  $g = \{g_i : 1 \leq i \leq n\}$ . Moreover, the numerical differentiation  $g'$  can be also represented by an affine transformation:

$$\{(g')_i : 1 \leq i \leq n\} = Cg + d \quad C \in \mathbb{R}^{n \times n}, \quad d \in \mathbb{R}^n$$

where the dimensions of  $A$ ,  $C$  and  $b$ ,  $d$  may vary depending on the numerical differentiation scheme. The discretized minimization problem becomes a quadratic program:

$$\min_g \alpha \delta t \cdot (Cg + d)^T (Cg + d) + \delta t \cdot (Ag + b - f)^T (Ag + b - f) \quad (4.40)$$

with some minor constraints whenever applicable (such as the non-negativity of  $v$  when differentiating  $x$ ), where  $\delta t$  denotes the time step size. This mathematical program can be solved efficiently with standard solvers given that the dimension of  $f$  or  $g$  is not too high (no more than a few thousands), which is indeed the case for the NGSIM dataset.

**Remark 4.2.** Other choices of the regularization term  $R(g)$  include the  $L^1$  norm of the derivative of  $g$ , i.e.

$$R(g) = \int_0^T |g'(\tau)| d\tau \quad (4.41)$$

In this case, due to the presence of the absolute value, the discrete optimization problem becomes non-smooth. Chartrand (2011) solves this case by applying first-order necessary conditions to the continuous-time problem (4.39), which are then discretized. Another choice for  $R(g)$  is the square of the  $L^2$  norm of  $g$  (Tikhonov, 1963):

$$R(g) = \int_0^T |g(\tau)|^2 d\tau \quad (4.42)$$

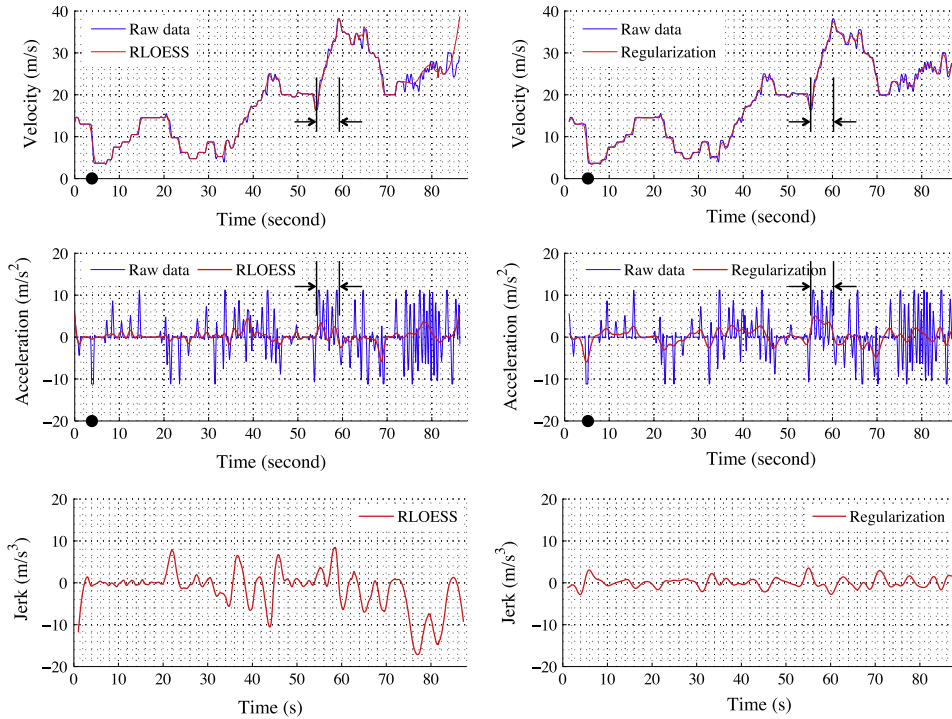
which suppresses large values of  $g$  without necessarily forcing it to be regular. A quadratic program can be formulated in this case by taking  $C$  to be the identity matrix and  $d = 0$  in (4.40).

#### 4.3. Implementation of the denoising methods

This section presents a numerical study of the two denoising methods mentioned above. We apply the two methods to the NGSIM dataset with the following chosen parameters. For the robust smoothing method, we apply the “RLOESS” to location (with span = 0.005), velocity (with span = 0.01) and acceleration (with span = 0.03) separately. We note that by doing so, the trajectories are no longer consistent; e.g. the derivative of the smoothed location is not approximated by the smoothed velocity. Regarding the robust differentiation method, we apply the regularization in differentiating location with  $\alpha = 0.01$  to get velocity, and then differentiate the acquired velocity with  $\alpha = 0.05$  to get the acceleration. The resulting trajectories are relatively consistent thanks to the data fidelity term in (4.39). Notice that in both methods, we have applied different degrees of smoothness (span in the first method and  $\alpha$  in the second method) to location, velocity and acceleration due to the natural degradation of regularity as a result of differentiation.

Fig. 1 compares the denoised velocity and acceleration profiles of the same car obtained by the two methods. A discernible difference in the acceleration can be seen while the velocity profiles are similar. We illustrate the qualitative difference

<sup>5</sup> Most numerical integration or differentiation operators are linear in the discrete values of the subject function, therefore they can be represented by affine transformations.



**Fig. 1.** Comparison of denoised velocity (top row), acceleration (middle row), and jerk (bottom row) obtained from RLOESS (left column) and regularization (right column). The bottom row only shows the denoised data as the raw data oscillate violently and obstruct the visualization.

between the two acceleration profiles at two distinct time windows. First, at  $t \approx 4$  s which is marked with a black dot in the figures, the vehicle experienced a sudden velocity reduction, indicating one or several large negative values in the acceleration. This is, however, insufficiently captured by RLOESS as these negative values were likely to be treated as outliers. The regularized differentiation, on the other hand, captures this negative acceleration. The second time window is roughly [55, 60] s, during which the car is continuously accelerating. This is reflected by the regularization method which predicts an acceleration well above zero during this period. However, the RLOESS method yields some negative acceleration values. Overall, it is confirmed that while both methods yield improved regularity in the estimated quantities, the regularization method is more faithful to the physical interpretation of these quantities, as expected from the theory. The RLOESS sometimes treats unusual yet realistic measurements as outliers (such as the sudden speed reduction at  $t = 4$  s). In addition, the regularization method yields acceleration that does not change signs as frequently as the RLOESS, which is consistent with the expected vehicle performance.

In the last row of Fig. 1, we show the time-trajectories of the jerk (rate of change of the acceleration) obtained from the two methods, with  $\text{span} = 0.05$  and  $\alpha = 0.1$ , respectively. An inspection of the overall estimation of velocity, acceleration and jerk reveals that the two methods yield estimation results that are increasingly different when the order of the underlying traffic quantity is higher, which makes sense as the noises are amplified and the regularities are worsened each time when differentiation is performed. In addition, the regularization method seems to represent the physics of traffic relatively well, while the RLOESS method may be more resilient to the outliers in the data and sensing errors.

For the study presented in the rest of this paper, we re-construct the ground truth using the following two-step method. We first apply the RLOESS method with  $\text{span} = 0.01$  to smooth the location data. Then, we use the smoothed location to perform robust numerical differentiations following Section 4.2, and the resulting velocity (with  $\alpha = 0.01$ ) and acceleration (with  $\alpha = 0.03$ ) are taken as the ground truth.

**Remark 4.3.** The proposed two-step method yields location  $x$  and acceleration  $a$  that are consistent in the sense of regularized differentiation, rather than finite-difference differentiation (i.e.  $a_i = \frac{x_{i+1} - 2x_i + x_{i-1}}{\Delta t^2}$ ). One could devise a further step to construct another location data  $\tilde{x}$  that is consistent with  $a$  in the finite-difference sense, although this step is not a necessity for our study. The way to do this is to treat acceleration  $a$  as the ground truth and construct  $\tilde{x}$  according to the following optimization problem

$$\tilde{x} = \underset{y}{\operatorname{argmin}} \|y - x\|^2 \quad \text{subject to} \quad \mathcal{D}_2 y = a$$

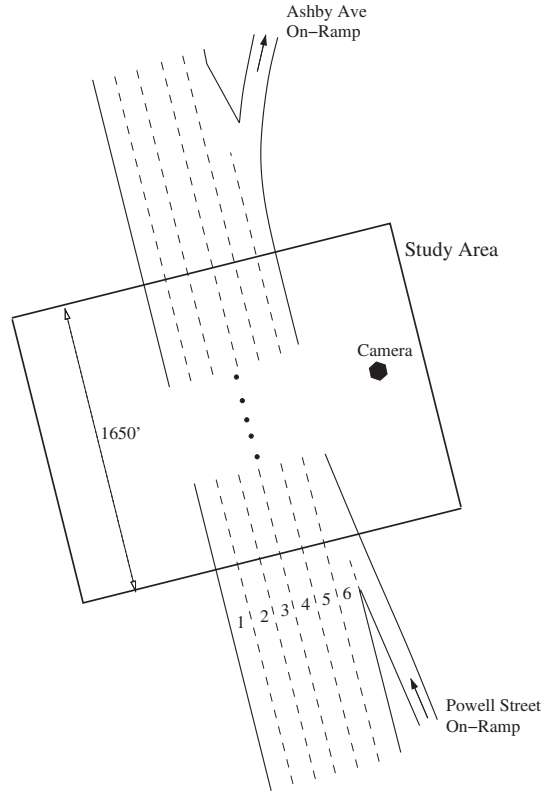


Fig. 2. The study area, spanning 1650 feet in length in the northbound of Interstate 80 located in Emeryville, CA.

where the matrix operation  $\mathcal{D}_2 y$  represents the second-order differentiation of  $y$  using finite difference (2.3), which is expressible as a linear operator. The constraint  $\mathcal{D}_2 y = a$  is the consistency condition, and the objective is a data fidelity condition that minimizes the discrepancy between  $\tilde{x}$  and the smoothed data  $x$ . This minimization problem can be formulated as a quadratic program.

Applying this method will yield a single trajectory  $\tilde{x}$ , whose second derivative (in the finite difference sense) is  $a$ , the smoothed acceleration from our two-step method.

## 5. Model calibration

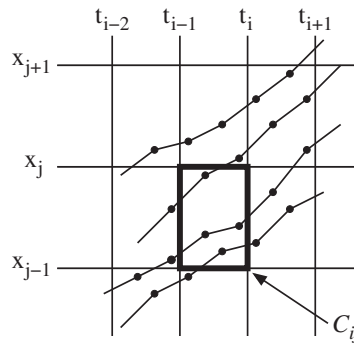
Before we implement the computational procedures proposed in Section 3, we need to estimate the values for  $A$ ,  $B$  and  $\rho_{jam}$  appearing in (3.10) and subsequent calculations. We use the available NGSIM dataset to calibrate these modeling parameters.

### 5.1. Description of vehicle trajectory data

The NGSIM program focuses on the northbound of I-80 located in Emeryville, CA. The highway segment of interest spans 1650 feet in length with an on-ramp at Powell Street and an off-ramp at Ashby Avenue; see Fig. 2. The highway segment has six lanes with the leftmost lane being a *high-occupancy vehicle* (HOV) lane, and the rightmost one being a merge/diverge lane. Data were collected using several video cameras. Digital video images were collected over an approximately five-hour period from 2:00 pm to 7:00 pm on April 13, 2005. Complete vehicle trajectories transcribed at a resolution of 1 frame per 0.1 s, along with vehicle type, lane identification and so forth were recorded and processed over three time slots: 4:00–4:15 pm, 5:00–5:15 pm, and 5:15–5:30 pm. The layout of the study area is shown in Fig. 2.

### 5.2. Estimating the density–flow relationship

We begin with estimating the density–flow relationship (3.8) needed for the congested phase represented by Eq. (3.6). Unlike the LWR model where the density–flow relation is expressed as a single-valued function, the fundamental diagram corresponding to the congested phase of the PTM is a multi-valued map. This means that a given density  $\rho$  corresponds to a



**Fig. 3.** Estimation of density and velocity inside a bin. The curves represent vehicle trajectories where the locations are recorded at the solid dots. For the depicted scenario, the occupancy of bin  $C_{ij}$  is three vehicles; the average velocity is taken as the mean of velocities measured/calculated at the four dots inside  $C_{ij}$ .

range of velocities  $v(\rho, q)$ ,  $q - q^* \in [-1, 1]$ . In order to identify this multi-valued function, we partition the temporal-spatial domain into small bins  $C_{ij} \triangleq [t_{i-1}, t_i] \times [x_{j-1}, x_j]$ ,  $i = 1, \dots, N_T, j = 1, \dots, N_X$  where  $i$  and  $j$  indicate the time step and the spatial step respectively. The average density associated with  $C_{ij}$  is estimated by the number of vehicles whose trajectories indicate their presence in the road segment  $[x_{j-1}, x_j]$  during time interval  $[t_{i-1}, t_i]$ . The velocity inside  $C_{ij}$  is calculated as the mean of all velocity measurements collected within this bin. Fig. 3 illustrates such a procedure. The dimension of the bins used to construct the density–flow relation in the congested phase is 4 (seconds)  $\times$  400 (feet).

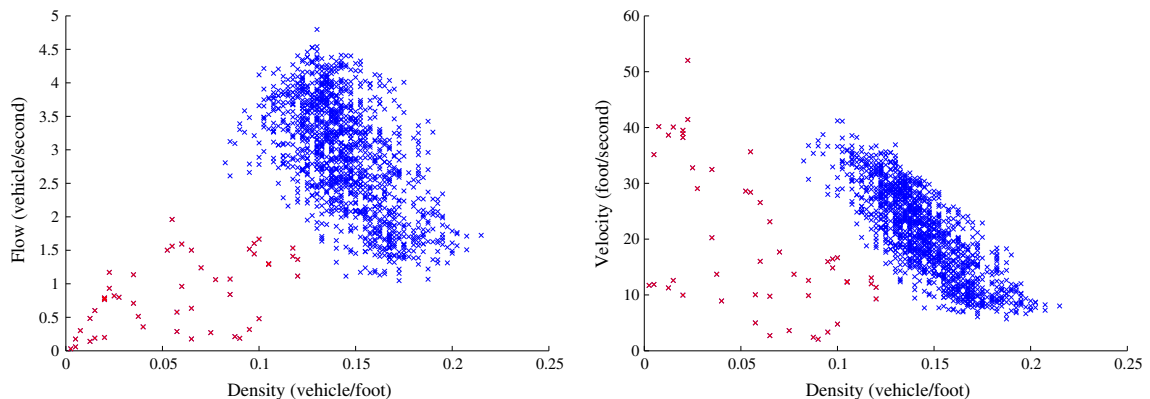
The flow within  $C_{ij}$  is calculated as the product of the density and the average velocity. The resulting density–flow and density–velocity plots for the entire study period are shown in Fig. 4.

### 5.3. Constructing the congested region in the fundamental diagram

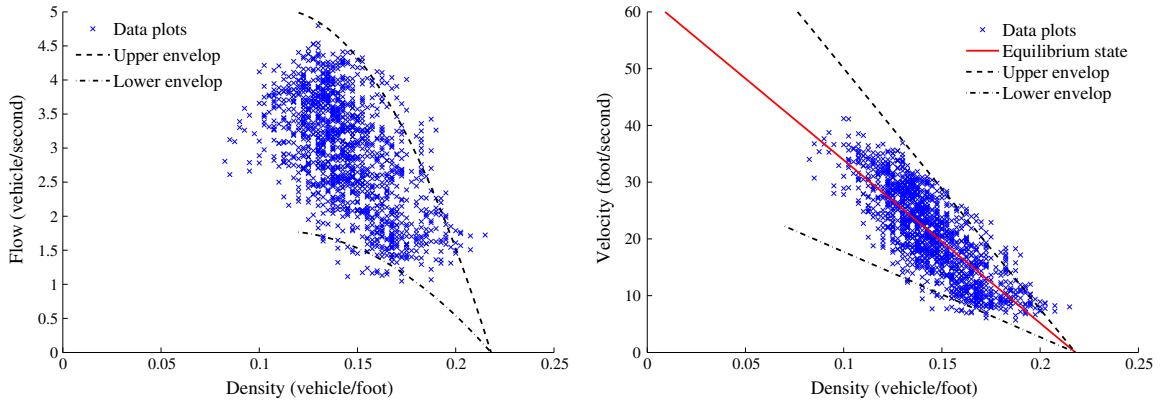
Eq. (3.10) suggests an affine density–velocity relationship  $v = A(\rho_{jam} - \rho)$  when the perturbation  $q - q^*$  is zero, a situation referred to as the *equilibrium state*. To determine  $A$  and  $\rho_{jam}$ , we conduct the following procedure. First, we visually remove the outliers in the density–velocity plot; see Fig. 4. Then, we conduct a linear regression to determine the equilibrium density–velocity relationship:  $v = A(\rho_{jam} - \rho)$ . The best fit is  $v = -287.25\rho + 62.61$ , which implies  $A = 287.25$  (foot<sup>2</sup>/vehicle/second) and  $\rho_{jam} = 0.218$  (vehicle/foot). The linear fit is indicated as the red solid line in the right part of Fig. 5.

It remains to determine  $B$ , which effectively dictates the width of the 2-D area in the congested domain of the fundamental diagram. Since  $q - q^* \in [-1, 1]$ , the upper and lower envelopes of the congested domain in the density–velocity relationship, depicted in the right part of Fig. 5, are respectively

$$\begin{cases} A(\rho_{jam} - \rho) + B(\rho_{jam} - \rho) & \text{(upper envelop)} \\ A(\rho_{jam} - \rho) - B(\rho_{jam} - \rho) & \text{(lower envelop)} \end{cases}$$



**Fig. 4.** The congested branch of the density–flow relationship (left) and the density–velocity relationship (right) for the PTM, both expressed as set-valued functions of density. The red crosses represent outliers that are discarded in the curve fitting. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** The fitted density–flow relationship (left) and density–velocity relationship (right) with  $A = 287.25$  (foot<sup>2</sup>/vehicle/second),  $B = 137$  (foot<sup>2</sup>/vehicle/second), and  $\rho_{jam} = 0.218$  (vehicle/foot).

Since  $A$  and  $\rho_{jam}$  are known, we choose the smallest value of  $B$  such that  $>95\%$  of the data points fall within the area formed by the upper and lower envelopes; see the right part of Fig. 5. This yields  $B = 137$  (foot<sup>2</sup>/vehicle/second), and the resulting envelopes are shown in the right part of Fig. 5. Consequently, the upper and lower envelopes in the density–flow relationship are uniquely determined as

$$\begin{cases} A(\rho_{jam} - \rho)\rho + B(\rho_{jam} - \rho)\rho & \text{(upper envelop)} \\ A(\rho_{jam} - \rho)\rho - B(\rho_{jam} - \rho)\rho & \text{(lower envelop)} \end{cases},$$

which are quadratic curves (parabolas); see the left part of Fig. 5.

## 6. Estimating traffic quantities along vehicle trajectories

In this section, we present the estimation results associated with various first- and higher-order traffic quantities along vehicle trajectories. Those quantities include: velocity and acceleration given by (2.2) and (2.3), vehicle density given by (3.14), (3.20) or Algorithm 1, deviation given by (3.15), (3.21) or Algorithm 1, and the power demand and HC emission rate given by (3.36) and (3.37). Different sampling frequencies will be used to investigate the deterioration of estimation quality caused by under sampling. Recall that the vehicle locations in the NGSIM raw dataset are recorded every 0.1 s. In order to accommodate different sampling frequencies, we consider an integer  $N$  and extract data from the raw dataset every  $N$  points. For example,  $N = 30$  implies a sampling period of  $0.1 \times 30 = 3$  s. We then compute the relative  $L^1$  error between the quantities estimated with  $N > 1$  and the ground truth.

All the numerical results reported below are based on data collected during 4:00–4:15 pm and 5:00–5:30 pm. The total number of vehicles involved in these time periods is 5677. The results summarized in tables or figures below are based on data in the entire study period of 45 min unless otherwise stated. As we have explained in detail in Section 4, the ground-truth quantities used throughout our numerical experiments are re-constructed using a combined approach of smoothing and robust differentiation.

### 6.1. Velocity and acceleration

Vehicle speed and acceleration are among the most fundamental traffic quantities. Following the steps explained at the beginning of this section, we compute and summarize in Table 1 the mean and standard deviation of the relative  $L^1$  errors for the velocity and acceleration. Overall, the estimation of both these quantities deteriorates with under sampling. Moreover, the acceleration estimation is more susceptible to under sampling than velocity estimation. This is because much of the

**Table 1**

Relative errors of velocity and acceleration estimation using different sampling frequencies.  $v_{true}$  and  $a_{true}$  denote the ground-truth velocity and acceleration respectively;  $v_N$  and  $a_N$  denote the reconstructed velocity and acceleration with different values of  $N$ .

$N$	Mean (%)			Standard deviation (%)		
	10	20	30	10	20	30
$\frac{\ v_{true} - v_N\ _{L^1}}{\ v_{true}\ _{L^1}}$	2.15	4.35	6.41	0.63	1.46	2.53
$\frac{\ a_{true} - a_N\ _{L^1}}{\ a_{true}\ _{L^1}}$	20.91	47.43	63.51	6.23	11.49	11.94

higher-order variation in the acceleration takes place on a smaller time scale, as confirmed by the poor regularity in the acceleration profile (see, for example, Figs. 1 and 6), and cannot be sufficiently captured with under sampling. Fig. 6 shows one example of the reconstructed velocity and acceleration of the same car by using different values of  $N$ . It is confirmed that the fitting of velocity shows an overall improvement over the fitting of acceleration, which is caused by the poorer regularity in the latter quantity, i.e. more oscillations are observed in a short time period.

## 6.2. Vehicle density and deviation

The density  $\hat{\rho}$  and deviation  $\hat{q}$  are estimated by (3.14) and (3.15) (Method 1), (3.20) and (3.21) (Method 2), or Algorithm 1 (Method 3) under various assumptions on the source term and traffic flow conditions. The means and standard deviations of the relative  $L^1$  errors are summarized in Table 2.

All three methods show qualitatively similar results: The estimation of the first-order quantity (density) is more accurate than the estimation of the second-order quantity (deviation) under the same condition. Moreover, both estimates deteriorate as the sampling frequency decreases. These results show some similarities to the estimation of velocity and acceleration (Section 6.1). Furthermore, method 2 shows slightly worsened yet satisfactory estimation for density  $\hat{\rho}$ . On the other hand, its estimation of  $\hat{q}$  suffers significantly with under sampling (all errors are greater than 100% and thus are not shown in the table). As we have explained in Eq. (3.22) that the finite-difference approximation of the quantity  $\frac{Dv}{Dt} - vv_x$  is zero. Even with modified finite-difference schemes such as (3.23) or robust differentiation schemes mentioned in Section 4, such a quantity is likely to be very small, causing Eq. (3.21) to yield  $\hat{q}$  with small magnitude and small  $L^1$  norm, and consequently, leading to large relative  $L^1$  errors. This shows that Method 2 is not robust enough in handling the deviation  $\hat{q}$ .

Interestingly, the estimation errors for  $\hat{q}$  in the no-source case (Method 3) are much lower than Method 1. This is due to the exact approach employed by Method 3 without simplification regarding the traffic dynamics, as opposed to the other two approaches. However, recall that Method 3 works only when  $v > \frac{4}{3}A\rho_{jam} \approx 27.83$  (foot/s). As shown by the right half of Fig. 4, the majority of the speed measurements fall below this value. In order to apply this method, we select vehicles whose trajectories indicate at least one non-trivial period (over 20 s), during which the speed is uniformly above 27.83 foot/s, and then apply Algorithm 1 to these periods. The total number of vehicles involved in this computation is 1628 (out of 5677 in total). We conclude that the applicability of Algorithm 1 for the no-source case is limited, although it yields a more accurate estimation of the deviation than the other two methods.

Fig. 7 shows one example of the reconstructed density and deviation profiles along a single vehicle trajectory, for three different sampling frequencies ( $N = 10, 20, 30$ ). First of all, a comparison of these cases leads to a similar conclusion regarding the deterioration of estimating first- and second-order quantities as before. Moreover, one can clearly observe, from both ground-truth and reconstructed  $\hat{q} = q - q^*$ , that the vehicle frequently experienced changes in the traffic conditions, between below-equilibrium speed ( $q - q^* < 0$ ) and above-equilibrium speed ( $q - q^* > 0$ ). Such a characteristic of congested, unstable traffic cannot be captured by first-order models such as the LWR model as it describes only the equilibrium speed as a function of density.

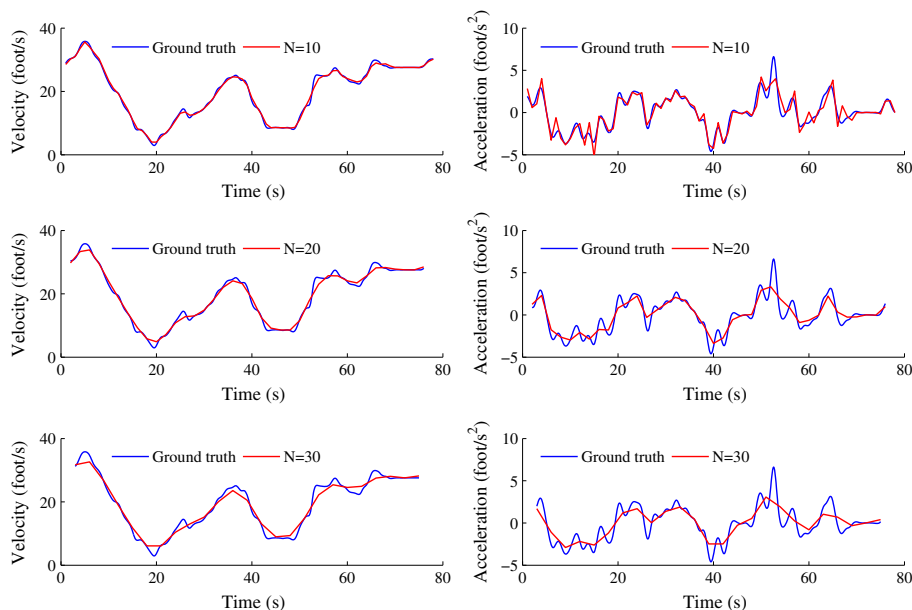


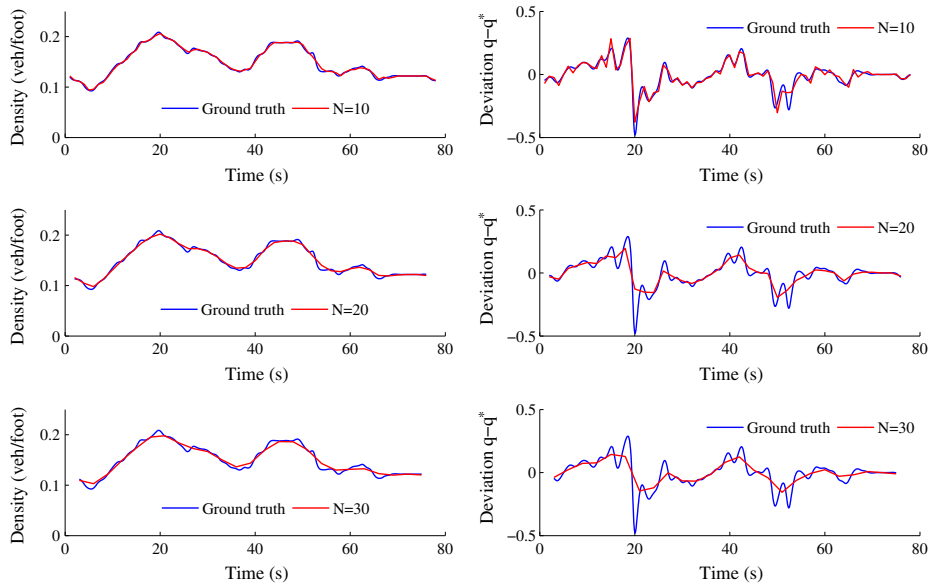
Fig. 6. The ground-truth vs. estimated velocity (left column) and acceleration (right column) of the same vehicle, using different sampling frequencies ( $N = 10, 20, 30$ ).



**Table 2**

Estimation of  $\hat{\rho}$  and  $\hat{q}$  based on the phase transition model with Methods 1–3.  $\hat{\rho}_{true}$  and  $\hat{q}_{true}$  represent the ground-truth quantities estimated using the ground-truth location, velocity and acceleration re-constructed in Section 4.  $\hat{\rho}_N$  and  $\hat{q}_N$  denote the estimated quantities with under sampling. The estimation of  $\hat{q}$  with Method 2 is not shown as the errors are above 100%.

	$N$	Mean (%)			Standard deviation (%)		
		10	20	30	10	20	30
$\frac{\ \hat{\rho}_{true} - \hat{\rho}_N\ _{L^1}}{\ \hat{\rho}_{true}\ _{L^1}}$	Method 1	2.15	4.55	6.64	0.63	1.62	2.70
	Method 2	3.56	6.70	9.35	1.35	2.73	4.08
	Method 3	7.19	13.80	19.39	7.48	14.37	19.46
$\frac{\ \hat{q}_{true} - \hat{q}_N\ _{L^1}}{\ \hat{q}_{true}\ _{L^1}}$	Method 1	20.91	54.41	69.32	6.23	10.96	10.23
	Method 2	–	–	–	–	–	–
	Method 3	12.80	27.58	40.75	6.30	13.43	19.22



**Fig. 7.** The ground-truth and estimated density (left column) and deviation (right column) profiles along the same vehicle trajectory, using different sampling frequencies.

**Table 3**

Estimations of the power demand functions  $Z$  and the hydrocarbon emission rate  $r^{HC}$  based on vehicle trajectories. Subscripts *true* and *N* denote the ground-truth quantity and the reconstructed quantity, respectively.

$N$	Mean (%)			Standard deviation (%)		
	10	20	30	10	20	30
$\frac{\ r_{true}^{HC} - r_N^{HC}\ _{L^1}}{\ r_{true}^{HC}\ _{L^1}}$	13.39	23.58	27.30	16.77	24.69	24.70
$\frac{\ Z_{true} - Z_N\ _{L^1}}{\ Z_{true}\ _{L^1}}$	18.69	44.93	60.42	4.95	9.36	10.47

### 6.3. Emission rate and power demand function

The HC emission rate and the power demand function are estimated according to (3.37) and (3.36) respectively, where the velocity  $v$  and acceleration  $a$  are approximated by (2.2) and (2.3). Table 3 summarizes the relative  $L^1$  errors in estimating the HC emission rate  $r^{HC}$  and the power demand  $Z$ , when different sampling frequencies are employed.

Overall, estimation of these higher-order quantities suffer from under sampling since higher-order variations in the acceleration profiles are ignored, as expected from Section 6.1. Moreover, the deterioration in the accuracy seems most significant when the sampling period increases from  $N = 10$  (1 s) to  $N = 20$  (2 s), which suggests that the sampling period should ideally

be well below 2 s in order to capture these quantities. This can be confirmed from Fig. 8, which shows both ground-truth and estimated  $r^{HC}$  and  $Z$  associated with the same vehicle. We see that while  $N = 10$  captures the variations in emission rate and power demand relatively well,  $N = 20$  or 30 causes the estimation to miss the majority of the peak values. Finally, a comparison between  $r^{HC}$  and  $Z$  for the same value of  $N$  reveals that the errors in the former case is much smaller. The reason is that while both estimates misinterpret the high variations by similar amount, the time-trajectory of  $r^{HC}$  is uniformly above some positive value while the time-trajectory of  $Z$  remains centered at zero. Therefore, the relative  $L^1$  error for  $r^{HC}$  tends to be smaller.

## 7. Estimating Eulerian quantities

The previous section is mainly concerned with traffic quantities associated with a Lagrangian particle; i.e. a moving car. It would be desirable to further explore the effect of under sampling in an Eulerian framework; that is, we will study traffic quantities represented on the spatial–temporal domain. We are also prompted to examine how the estimating error depends on the penetration rate of probe vehicles (mobile sensors), in addition to the sampling frequency. The Eulerian traffic quantities studied in this section include vehicle density and deviation. Vehicle emission rate will be studied in the next Section.

### 7.1. Estimating vehicle densities

The procedure for estimating Eulerian density based on vehicle trajectory data and the phase transition model is illustrated as follows. We consider again the temporal-spatial bins  $C_{ij}$ ,  $i = 1, \dots, N_T$ ,  $j = 1, \dots, N_X$ , each expressed as a product of intervals  $[t_{i-1}, t_i] \times [x_{j-1}, x_j]$ . Recall that given a discrete-time trajectory of a vehicle:

$$\dots, x(\tau_{k-1}), x(\tau_k), x(\tau_{k+1}), \dots$$

where  $\{\tau_k\}$  is a fixed time grid, one can estimate the velocity  $v(\tau_k)$ , acceleration  $a(\tau_k)$ , and the Lagrangian density  $\rho(\tau_k, x(\tau_k))$  using techniques elaborated in Sections 2 and 3. In order to estimate the average density of  $C_{ij}$ , we search for the probe vehicles whose trajectories intersect  $C_{ij}$ . Then the density  $\rho_{ij}$  associated with  $C_{ij}$  is estimated as the average of  $\rho(\tau_k, x(\tau_k))$  for all  $k$  such that  $(\tau_k, x(\tau_k)) \in [t_{i-1}, t_i] \times [x_{j-1}, x_j]$ .

One example of such a calculation is presented in Fig. 9, where we utilize vehicle trajectories with a sampling period of  $N = 30$  (3 s) and a 50% probe penetration rate to perform the estimation of bin-based densities, which is then compared with the densities obtained from  $N = 1$  and a 100% penetration rate. Fig. 9 shows consistency between the two density estimations. Notice from the bottom of Fig. 9 that a few bins are marked with dark blue, which means that they do not have any valid location measurements that support the computation of density values. We call these bins *inactive*, which are automatically assigned zero density and indicated by dark blue in Fig. 9. It is expected that as the penetration rate and the sampling frequency decrease further, the number of inactive bins will increase drastically. To avoid this, one should choose the dimension of a bin reasonably large to ensure having at least some valid measurements inside each bin. In

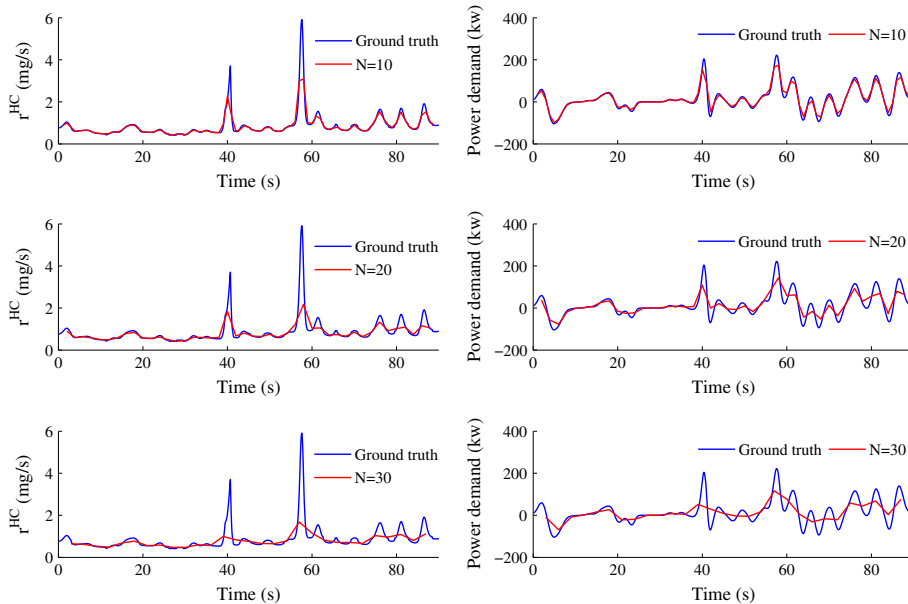


Fig. 8. Ground-truth and estimated hydrocarbon emission rate (left column) and power demand function (right column).

our numerical results presented in the next section, the dimension of a bin is chosen to be 4 (seconds)  $\times$  400 (feet). We shall use the phase transition model and the LWR model to perform Eulerian density estimation with a range of sampling frequencies and probe penetration rates.

## 7.2. Density estimation: a comparative study of the PTM and the LWR model

In this section we evaluate the performance of the aforementioned estimation method in the presence of insufficient data coverage, that is, when the penetration rate and the sampling frequency are low. The purpose is to evaluate and justify the effectiveness of the proposed computational procedure, and to identify certain range of penetration rates and sampling frequencies such that the domain of study is sufficiently covered and the estimation error remains reasonably low.

In addition, we propose a comparative study of second-order models (the phase transition model) and first-order models (the LWR model) in terms of their accuracies in density estimation. For the LWR model (3.5), we need to calibrate the single-valued fundamental diagram (FD) in view of the density–flow relationship shown in Fig. 4. To do this, we first consider an affine congested branch by applying a least-square linear regression to the density–flow plot after removing the outliers; see the left figure of Fig. 10. The result, however, was rather peculiar (with very large  $\rho_{jam}$ ) as shown in Fig. 10. Applying this affine FD leads to substantial error in the estimation and thus is not adopted here. We instead employ the equilibrium density–flow relationship from the PTM case (notice that this relationship is obtained from a least-square linear regression of the density–velocity plots; see the right part of Fig. 10):

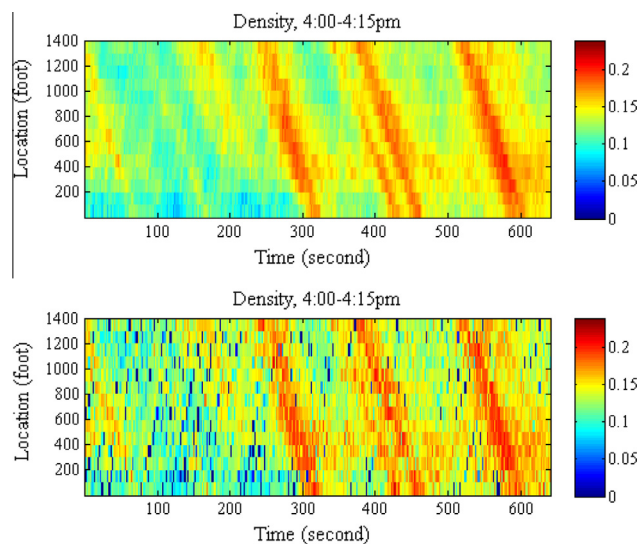
$$\text{flow} = A\rho(\rho_{jam} - \rho)$$

where  $A = 287.25$  (foot<sup>2</sup>/vehicle/s),  $\rho_{jam} = 0.218$  (vehicle/foot); see Fig. 10 for the curve fitting result, where the utilized curves are indicated with solid red lines.

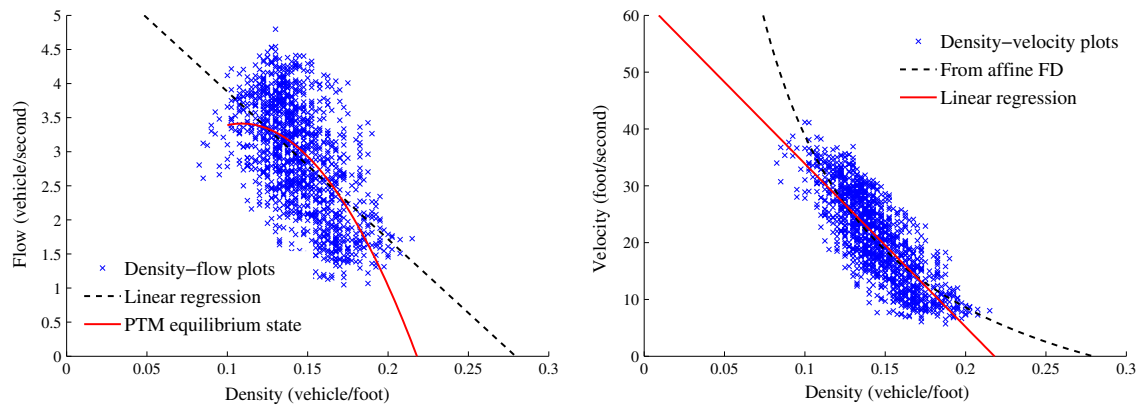
It can be seen that the single-valued density–flow relationship, no matter how derived, does not capture deviation from the equilibrium states, i.e. the scattered measurements in a 2-D region. The LWR model treats the traffic stream as stable and homogeneous; as a result, second-order variations such as acceleration/deceleration are not sufficiently captured. In order to apply the LWR model to perform density estimation, we first utilize the probe trajectory data to obtain information of the instantaneous speed  $v$ ; then we evaluate the function  $\rho = g^{-1}(v)$  to get the estimated density value, where  $g(\rho)$  is the invertible density–speed relationship depicted on the right half of Fig. 10. The rest of the procedure is the same as the PTM case.

The ground-truth density associated with each spatial–temporal bin is calculated by counting the number of vehicles present in that bin, using ground-truth trajectory data. The ground-truth density is then compared with the estimated densities obtained with a combination of lower sampling frequencies and lower penetration rates. The relative errors for all active bins are averaged and shown in Table 4 for period 4:00–4:15 pm, in Table 5 for period 5:00–5:15 pm, and in Table 6 for period 5:15–5:30 pm. Here in these calculations, Method 1 is employed by the PTM-based estimation. The results with Method 2 is qualitatively similar and are omitted here.

From all three tables, we notice that the penetration rate of probe vehicles has a stronger effect on the estimation accuracy than the sampling frequency. This again confirms our observation that, when it comes to estimating first-order



**Fig. 9.** Reconstruction of cell densities (in vehicle/foot) during 4:00–4:15 pm. The top figure shows the density obtained with  $N = 1$  and 100% penetration rate. The bottom figure shows the estimated density with  $N = 30$  and penetration rate = 50%. The inactive bins are assigned zero density and marked as dark blue in the figure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 10.** Fitting of the congested branch of the LWR fundamental diagram. Our study uses the equilibrium density–velocity or density–flow relationship from the PTM case, which are indicated as solid red lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

quantities, under sampling does not affect the accuracy as much as the penetration rate. On the other hand, the substantial variation in the penetration rate, from 100% to 2%, has caused approximately 14% difference in both PTM-based and LWR-based estimations. Moreover, the PTM-based estimation is more accurate than the LWR-based estimation for any given combination of penetration rate and sampling frequency, and the difference is more evident for lower penetration rates. This is explained by the fact that the LWR model only captures the equilibrium state of the density–flow or density–velocity relationship, and ignores possible deviations from such an equilibrium state. As shown in Fig. 10, the congested traffic is unstable and likely to deviate from the equilibrium state. Thus, the second-order model (PTM) is able to capture these deviations through the use of the variable  $q$  and a multi-valued fundamental diagram. As a result, the second-order model is able to estimate vehicle densities more accurately than the LWR model for congested highway traffic.

### 7.3. Deviation of traffic from equilibrium

In this section we calculate  $q - q^*$ , the indicator of the deviation of a vehicle speed from the equilibrium speed of surrounding vehicles. The PTM with a multi-valued FD allows disequibrated traffic and complex waves, commonly observed in congested highway traffic, to be described and modeled. This is in contrast to the LWR model which captures only the equilibrium traffic states.

Using a similar procedure described in Section 7.1, we compute the quantity  $\hat{q} = q - q^*$  using formula (3.15). It has been shown in Table 2 that the estimation of  $\hat{q}$  is very susceptible to under sampling and, potentially, lower probe penetration as well, we utilize the entire ground-truth data ( $N = 1$ , penetration rate = 100%) to reconstruct the bin-based  $\hat{q}$ , following a similar procedure described in Section 7.1. The result is shown in Fig. 11 for the traffic during 5:00–5:30 pm. It can be seen from the top figure showing the density that the traffic was significantly congested and multiple shock waves were present and propagated with positive or negative speeds. The picture in the middle shows the bin-based  $q - q^*$ , which takes value in  $[-1, 1]$ . We observe that traffic stayed close to equilibrium ( $q - q^* = 0$ ) where density was relatively low (typically below 0.16 vehicle/foot). However, significant deviation occurred with higher traffic densities. Furthermore, the presence of these high densities was accompanied by a mixture of above-equilibrium ( $q - q^* > 0$ ) and below-equilibrium ( $q - q^* < 0$ ) states.

**Table 4**

Comparison of the PTM and LWR models in estimating bin-based densities using different sampling frequencies and penetration rates, for the period 4:00–4:15 pm. “Coverage rate” is the ratio of the number of active bins and the total number of bins.

		Sampling period	Probe vehicle penetration rate					
			100 %	50 %	20 %	10 %	5 %	2 %
PTM	Average error (%)	$N = 10$	8.74	9.25	11.04	13.40	16.95	21.53
		$N = 20$	8.86	9.44	11.19	13.48	17.05	21.26
		$N = 30$	8.99	9.57	11.41	13.74	17.25	21.04
	Coverage rate (%)		100.00	100.00	100.00	99.83	97.67	79.70
LWR	Average error (%)	$N = 10$	12.66	13.29	15.47	18.37	22.45	27.52
		$N = 20$	12.58	13.24	15.41	18.29	22.33	27.04
		$N = 30$	12.43	13.03	15.31	18.26	22.26	26.44
	Coverage rate (%)		100.00	100.00	100.00	99.83	97.67	79.70

**Table 5**

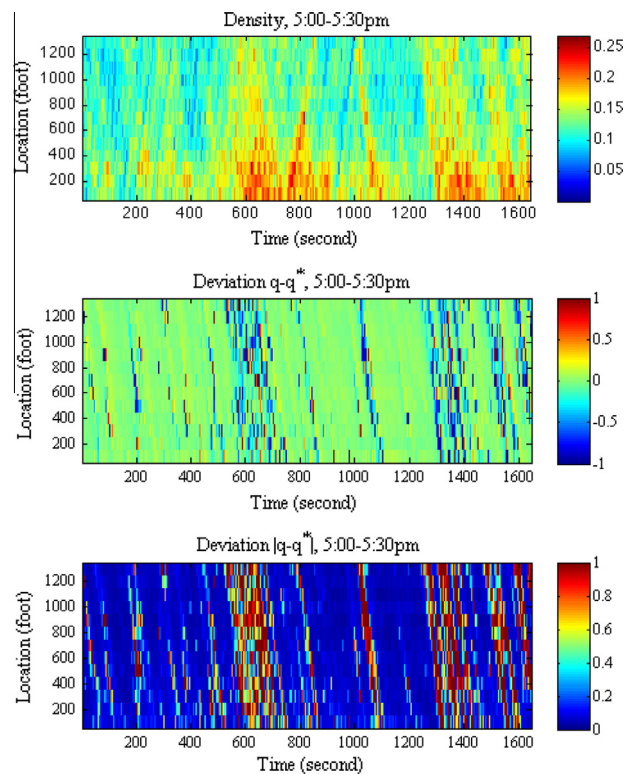
Comparison of the PTM and LWR models in estimating bin-based densities using different sampling frequencies and penetration rates, for the period 5:00–5:15 pm. “Coverage rate” is the ratio of the number of active bins and the total number of bins.

			Probe vehicle penetration rate					
			100 %	50 %	20 %	10 %	5 %	2 %
PTM	Average error (%)	$N = 10$	8.38	8.78	9.60	11.27	14.19	17.92
		$N = 20$	8.49	8.90	9.71	11.38	14.21	17.83
		$N = 30$	8.85	9.23	10.01	11.63	14.47	17.76
	Coverage rate (%)		100.00	100.00	100.00	99.97	98.57	82.88
LWR	Average error (%)	$N = 10$	10.78	11.26	12.24	14.24	17.85	23.09
		$N = 20$	10.96	11.44	12.44	14.44	17.91	22.95
		$N = 30$	11.44	11.93	12.93	14.81	18.29	22.80
	Coverage rate (%)		100.00	100.00	100.00	99.97	98.57	82.88

**Table 6**

Comparison of the PTM and LWR models in estimating bin-based densities using different sampling frequencies and penetration rates, for the period 5:15–5:30 pm. “Coverage rate” is the ratio of the number of active bins and the total number of bins.

			Probe vehicle penetration rate					
			100 %	50 %	20 %	10 %	5 %	2 %
PTM	Average error (%)	$N = 10$	8.89	9.00	9.62	10.42	12.24	15.60
		$N = 20$	9.02	9.13	9.74	10.53	12.35	15.47
		$N = 30$	9.17	9.29	9.93	10.71	12.52	15.39
	Coverage rate (%)		100.00	100.00	100.00	99.98	98.99	85.56
LWR	Average error (%)	$N = 10$	12.69	12.87	13.49	14.45	16.82	21.03
		$N = 20$	12.85	13.02	13.64	14.59	16.95	20.85
		$N = 30$	13.01	13.17	13.84	14.77	17.11	20.72
	Coverage rate (%)		100.00	100.00	100.00	99.98	98.99	85.56



**Fig. 11.** Space–time diagram of the ground-truth vehicle density (top), signed deviation  $q - q^*$  (center), and unsigned deviation  $|q - q^*|$  (bottom). The deviations are calculated according to (3.15).

We note that inside the same bin, above-equilibrium and below-equilibrium traffic may be present simultaneously; therefore, averaging  $q - q^*$  within the same bin does not fully capture the deviation. Thus, we further compute the bin-based quantity  $|q - q^*| \in [0, 1]$ , and the result is shown in the bottom picture of Fig. 11. Similar to the signed deviation  $q - q^*$ , the unsigned deviation  $|q - q^*|$  suggests a strong correlation with the vehicle density. This is consistent with the established observation that congested highway traffic tends to be unstable, due to inhomogeneous driving behavior (Laval and Leclercq, 2010) and the propagation of disturbances (Blandin et al., 2012).

## 8. Estimating instantaneous emission rate

It has been shown in Section 6.3 that the HC emission rate along vehicle trajectories are largely deteriorated by under sampling. Such significant error is caused by the high variation in the acceleration profile. However, it is expected that when aggregated on a cell or link level, the estimation accuracy of these quantities may improve even with large sampling periods. To confirm such an observation, we will next perform Eulerian-based HC emission estimation.

### 8.1. Bin-based estimation with insufficient data coverage

The bin-based emission rate can be estimated as follows. We first calculate the density  $\rho_{ij}$  for bin  $C_{ij}$  following the procedure illustrated in Section 7.1 using Method 1. Secondly, we use available vehicle trajectories to compute  $r^{HC}(\tau_k, x(\tau_k))$  using formula (3.37). Then all  $r^{HC}(\tau_k, x(\tau_k))$  that fall within  $C_{ij}$  are collected and averaged, which give rise to the average instantaneous HC emission rate per vehicle in this bin. Finally, this rate is multiplied by the number of vehicles present in this bin, which is given by the density, to yield the HC emission rate of  $C_{ij}$ .

Following this procedure, we perform the Eulerian estimation of  $r^{HC}$  for a combination of lower sampling frequencies and penetration rates, which are then compared with the ground truth. The bin-based relative errors are averaged and shown in Table 7 for the period 5:00–5:30 pm. Compared to the Lagrangian (single-vehicle) estimation, the Eulerian estimation shows improved accuracy due to the effect of spatial-temporal aggregation; the reader is referred to Table 3 for a comparison of the results. Interestingly, the influence of under sampling is reduced as the penetration rate becomes lower. This may be explained from the following two potential sources of error:

1. the lower sampling frequency that causes high variations in the emission rate to be ignored;
2. the computational procedure which directly combines a microscopic emission model with a macroscopic traffic model, and creates certain mismatch in these very different modeling scales.

It is likely that with very high penetration rate (e.g. 100%, 50%), the first factor is the dominating one, while for lower penetration rates the second factor (i.e. representing the emission rate inside a bin using a few microscopic vehicle trajectories) becomes the major source of error.

### 8.2. Correction factor approach for estimating aggregate emission rate

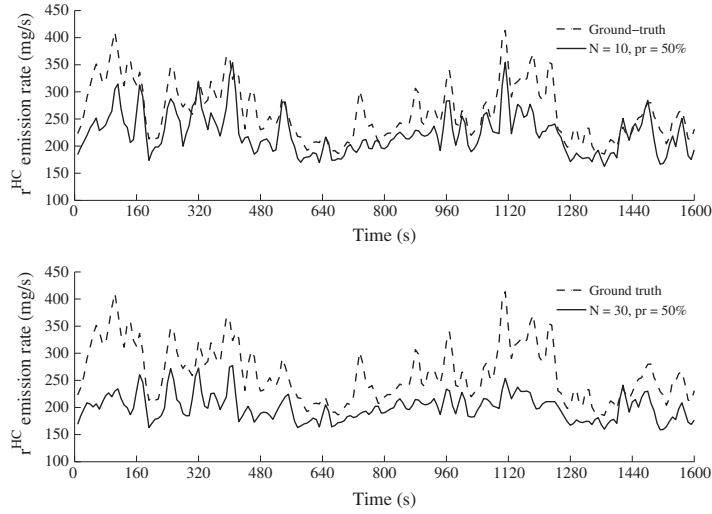
To further aggregate the hydrocarbon emission rate, we compute the total emission rate on the entire road segment. This is done by summing up bin-based rates over all spatial cells for each time interval. An example of these time-dependent rates are shown in Fig. 12, where the ground-truth  $r^{HC}$  is compared with the estimated ones with  $[N = 10, pr = 50\%]$  and  $[N = 30 \text{ and } pr = 50\%]$  respectively. From these figures we see that the estimation captures, to some degree, the overall trends of the emission rate, although it tends to underestimate the true value. The reason is that the under sampling overlooks higher-order variations in velocity and acceleration, and thus does not capture surges in the emission rate; a good example can be seen in Fig. 8. And, such an underestimation is expected to become worse for larger sampling periods, as can be seen from a comparison between the two pictures in Fig. 12.

The numerical result shown in Fig. 12 reveals discernible errors in emission estimation, although the overall time-varying trends of  $r^{HC}$  is captured using our method. As we argued in Section 8.1, these errors are attributed to two factors, and to

**Table 7**

Estimation of bin-based hydrocarbon emission rate  $r^{HC}$  using different sampling periods and penetration rates. The time period is 5:00–5:30 pm. The estimation of vehicle density employs Method 1.

	Sampling period	Probe vehicle penetration rate			
		100 %	50 %	20 %	10 %
$r^{HC}$ Average error (%)	$N = 10$	11.04	14.09	19.36	23.67
	$N = 20$	16.66	17.96	20.52	23.21
	$N = 30$	19.55	20.40	21.75	23.20
Coverage rate (%)		100.00	100.00	100.00	99.88



**Fig. 12.** Time-dependent HC emission rate on the entire study area (1600 feet in length) for the time period 5:00–5:30 pm. The top figure employs  $N = 10$  and 50% penetration rate; the bottom figure employs  $N = 30$  and 50% penetration rate.

resolve these two issues simultaneously, we propose a correction factor approach to adjust the estimation process, which is demonstrably effective.

#### 8.2.1. Correction factors based on regression models

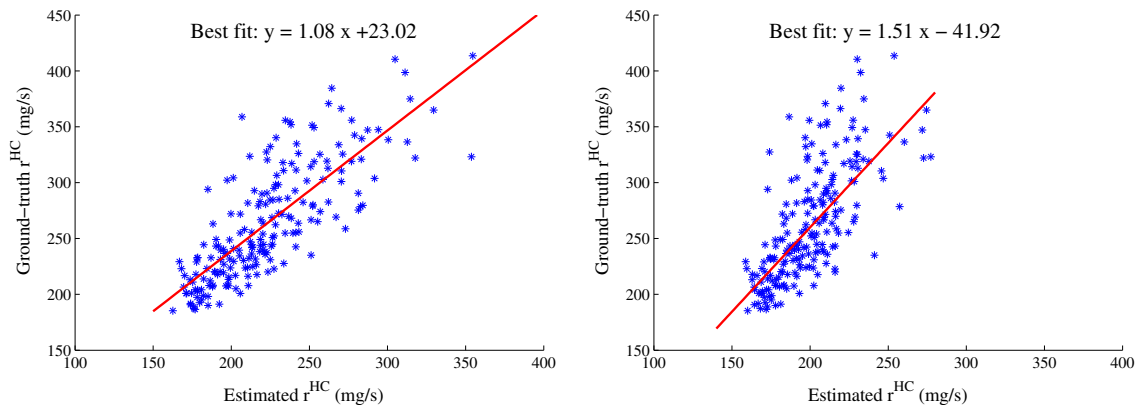
We employ a linear regression approach which finds an appropriate affine relationship between the ground-truth emission rates and the estimated ones. To demonstrated the hypothesized affine relationship, we plot in Fig. 13 the ground-truth value vs. the estimated value for the two cases illustrated in Fig. 12. Each data point on the plane corresponds to one time interval; the x-coordinate is the estimated value, and the y-coordinate is the true value. It is generally expected that different sampling periods and penetration rates require different linear regression coefficients. In fact, the best fit in the two cases are:

$$y = 1.08x + 23.02 \quad (N = 10, \text{ pr} = 50\%) \quad (8.43)$$

$$y = 1.51x - 41.92 \quad (N = 30, \text{ pr} = 50\%) \quad (8.44)$$

Indeed, the  $N = 30$  case underestimates emission rates more than the  $N = 10$  case, and thus has a larger 1st-order coefficient.

We denote by  $r_{true}^{HC}$  the ground truth emission rate, and by  $r_{est}^{HC}$  the estimated one. The linear model assumes that  $r_{true}^{HC} = \beta_0 + \beta_1 r_{est}^{HC}$  where  $\beta_0, \beta_1 \in \mathbb{R}$ . Applying the affine adjustment (8.43) and (8.44) yields the improved estimation depicted in Fig. 14. To rigorously justify the validity and robustness of the proposed correction factor approach, we employ a  $k$ -fold cross validation discussed below.



**Fig. 13.** Scatter plots of the estimated  $r_{est}^{HC}$  vs. ground-truth  $r_{true}^{HC}$  for the entire study area (1600 feet in length) during period 5:00–5:30 pm. The best linear fits are shown as solid lines. Left:  $N = 10, \text{ pr} = 50\%$ ; right:  $N = 30, \text{ pr} = 50\%$ .



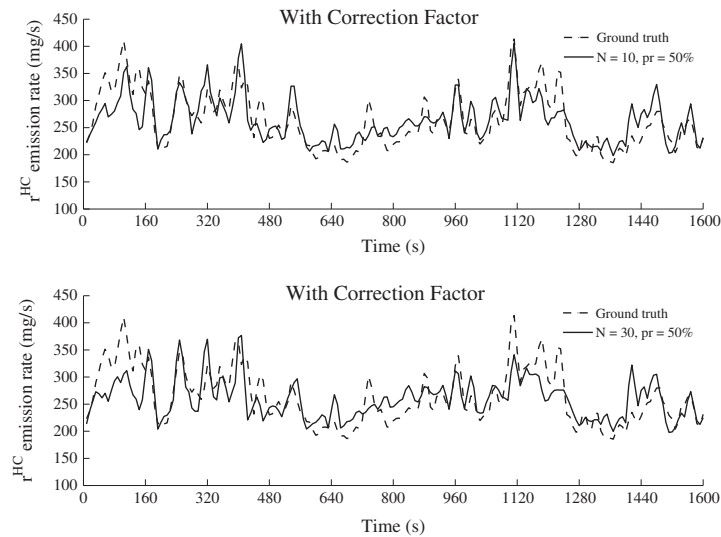


Fig. 14. The adjusted time series of  $r^{HC}$  with the correction factors (8.43) and (8.44).

### 8.2.2. Cross validation

We partition available NGSIM vehicle trajectories into  $k$  equal-size subsets. Among the  $k$  subsets, a single subset is used as the training data to obtain the linear coefficients  $\beta_0$  and  $\beta_1$ . The rest of the  $k - 1$  subsets are used to validate this affine adjustment by calculating the discrepancy between the predicted value (after applying the linear transformation) and the ground-truth value. Such a procedure is repeated  $k$  times, each with a distinct training dataset, and the errors from all  $k$  repetitions are averaged. Such a process is known as  $k$ -fold cross validation.

We use the absolute relative error  $|r_{est}^{HC}(t) - r_{true}^{HC}(t)|/|r_{true}^{HC}(t)|$  to indicate the accuracy of the estimation. A range of sampling periods and penetration rates are considered as shown in Table 8, which summarizes results of the  $k$ -fold cross validation. Notice that for each penetration rate  $pr$ , the value  $k$  is naturally chosen as  $1/pr$ . It is quite clear from these cross validation results that applying the correction factor yields significantly improved estimation accuracy.

**Remark 8.1.** It should be noted that the linear coefficients  $\beta_0$  and  $\beta_1$  employed by the correction factor approach are determined separately for each combination of sampling period and penetration rate. The reason is that these factors are devised to counter the effects of (1) underestimating the emission rates due to lower sampling frequency; and (2) possible mismatch between the microscopic emission model and macroscopic traffic model. Therefore, the factors must be determined separately as the effects (1) and (2) naturally vary in response to  $N$  and the penetration rate.

It can be seen that the estimation with correction factors produces errors that are not significantly affected by under sampling or lower penetration rate, with a maximum difference of 4% ( $[N = 10, pr = 50\%]$  vs.  $[N = 30, pr = 10\%]$ ) over all scenarios in Table 8. This is in contrast to the case without the correction factors, which has a degradation of accuracy by as large as 13% ( $[N = 10, pr = 50\%]$  vs.  $[N = 30, pr = 10\%]$ ). This is due to the fact that the correction factors have taken into account misinterpretations of density and emission rate caused by both larger  $N$  and smaller  $pr$ , and hence the estimation is more resilient to the change in both parameters.

## 9. Conclusion

This paper presents several approaches to estimate first- and higher-order traffic quantities based on vehicle trajectory data. It also conducts extensive study of the accuracy of these estimations with respect to the probe penetration rate and

Table 8

Cross validation results of the correction factor approach. The errors in the parenthesis are from calculations without any correction factor.

	Sampling period	Probe vehicle penetration rate		
		50%	20%	10%
$r^{HC}$ Average error (%)	$N = 10$	10.34 (11.24)	12.34 (15.60)	13.47 (19.77)
	$N = 20$	12.21 (17.57)	13.65 (19.77)	14.32 (22.17)
	$N = 30$	13.52 (21.89)	14.31 (23.17)	14.39 (24.22)

the sampling frequency of the mobile data. Most existing studies on mobile sensing tend to focus on the penetration rate, while the effect of sampling frequency is rarely studied. We present several data fusion schemes that incorporate vehicle trajectory data into the second-order phase transition model, and perform the estimation of first-order quantities such as speed and density, and higher-order quantities such as acceleration, deviation (perturbation) and emission rate. The main findings are summarized below.

- Various Eulerian and Lagrangian traffic quantities can be estimated via vehicle trajectory data, when the phase transition model (PTM) is employed.
- For Lagrangian estimation, the first-order quantities are accurately reconstructed by the proposed method, even with low sampling frequency. For higher-order Lagrangian quantities, lower sampling frequency significantly reduces the accuracy.
- The PTM-based estimation method is extended to perform Eulerian estimation, with varying penetration rate and sampling frequency. The estimation accuracy shows a trend similar to that of the Lagrangian estimation, and deteriorates with lower penetration rate.
- The PTM is compared with the LWR model in estimating Eulerian density with varying sampling frequency and penetration rate. The results show that the former yields more accurate results due to the fact that it captures second-order deviation of traffic, which is frequently observed in congested and unstable traffic and is not described by first-order models.
- A correction factor approach is applied to the Eulerian estimation of hydrocarbon emission rate, which significantly improves the accuracy. The method resolves the underestimation of emission due to insufficient sampling frequency, and the potentially biased emission estimation caused by mismatching modeling scales. As a result, the adjusted estimation is less sensitive to the changes in the sampling frequency and penetration rate.

Further research will focus on applying the methodology to urban arterial networks and comparing it with the LWR-based approach. As we mentioned earlier in the introduction, the benefit of applying the PTM to arterial traffic will be somehow limited. However, higher-order variations in acceleration, emission and fuel consumption still need to be captured and quantified in connection with a given traffic flow model, and the effects of sampling frequency and penetration rate on these estimations should be properly understood in an urban environment.

## Acknowledgement

The authors would like to thank two anonymous reviewers for their constructive remarks.

## References

- Ahn, K., Rakha, H., Trani, A., Van Aerde, M., 2002. Estimating vehicle fuel consumption and emission based on instantaneous speed and acceleration levels. *J. Transp. Eng.* 128, 182–190.
- Argote, J., Christofa, E., Xuan, Y., Skabardonis, A., 2011. Estimation of measures of effectiveness based on Connected Vehicle data. In: *Proceedings of the 14th International IEEE Conference on Intelligent Transportation Systems*, pp. 1767–1772.
- Aw, A., Rascle, M., 2000. Resurrection of “second order” models of traffic flow. *SIAM J. Appl. Math.* 60, 916–938.
- Ban, X., Herring, R., Hao, P., Bayen, A., 2009. Delay pattern estimation for signalized intersections using sampled travel times. *Transp. Res. Rec.* 2130, 109–119.
- Ban, X., Hao, P., Sun, Z., 2011. Real time queue length estimation for signalized intersections using sample travel times from mobile sensors. *Transp. Res. Part C* 19 (6), 1133–1156.
- Barth, M., An, F., Norbeck, J., Ross, M., 2007. Modal emission modeling: a physical approach. *Transp. Res. Rec.* 1520, 81–88.
- Blandin, S., Work, D., Goatin, P., Piccoli, B., Bayen, A., 2011. A general phase transition model for vehicular traffic. *SIAM J. Appl. Math.* 71, 107–127.
- Blandin, S., Argote, J., Bayen, A., Work, D., 2012. A phase transition model of non stationary traffic flow: definition, properties and solution method. *Transp. Res. Part B* 52, 31–55.
- Blatná, D., 2005. Robust regression. In: *Applications of Mathematics and Statistics in Economy*. Wroclaw.
- Chan, C.C., 1997. An overview of power electronics in electric vehicles. *IEEE Trans. Ind. Electron.* 44 (1), 3–13.
- Chartrand, R., 2011. Numerical differentiation of noisy, nonsmooth data. *ISRN Appl. Math.* 2011. <http://dx.doi.org/10.5402/2011/164564>. Article ID 164564.
- Cheng, Y., Qin, X., Jin, J., Ran, B., 2012. An exploratory shockwave approach to estimating queue length using probe trajectories. *J. Intell. Transp. Syst.* 16 (1), 12–23.
- Christiani, E., de Fabritius, C., Piccoli, B., 2010. Fluid dynamic approach for traffic forecast from mobile sensor data. *Commun. Appl. Ind. Math.* 1 (1), 54–71.
- Claudel, C.G., Bayen, A.M., 2011. Convex formulations of data assimilation problems for a class of Hamilton–Jacobi equations. *SIAM J. Control Optim.* 49 (2), 383–402.
- Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. *J. Am. Stat. Assoc.* 83, 596–610.
- Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I., 1990. STL: a seasonal-trend decomposition procedure based on loess. *J. Official Stat.* 6, 3–33.
- Colombo, R.M., 2002. Hyperbolic phase transitions in traffic flow. *SIAM J. Appl. Math.* 63 (2), 708–721.
- Colombo, R.M., Corli, A., 2002. Sonic hyperbolic phase transition and Chapman–Jouguet detonations. *J. Differ. Equ.* 184, 321–347.
- Colombo, R.M., Goatin, P., Priuli, F.S., 2007. Global well posedness of traffic flow models with phase transitions. *Nonlinear Anal.* 66, 2413–2426.
- Comert, G., Cetin, M., 2009. Queue length estimation from probe vehicle location and the impacts of sample size. *Eur. J. Oper. Res.* 197 (1), 196–202.
- Daganzo, C.F., 1994. The cell transmission model. Part I: a simple dynamic representation of highway traffic. *Transp. Res. Part B* 28 (4), 269–287.
- Daganzo, C.F., 1995. Requiem for second-order fluid approximations of traffic flow. *Transp. Res. Part B* 29 (4), 277–286.
- Demers, A., List, G.F., Wallace, W.A., Lee, E.E., Wojtowicz, J.M., 2006. Probes as path seekers: a new paradigm. *Transp. Res. Rec.: J. Transp. Res. Board* 1944, 107–114.
- Garavello, M., Piccoli, B., 2009. On fluidodynamic models for urban traffic. *Networks Heterogen. Media* 4 (1), 107–126.
- Goatin, P., 2006. The Aw–Rascle vehicular traffic flow model with phase transitions. *Math. Comput.* 44, 287–303.
- Herrera, J.C., Work, D.B., Herring, R., Ban, X.J., Jacobson, Q., Bayen, A., 2010. Evaluation of traffic data obtained via GPS-enabled mobile phones: the mobile century field experiment. *Transp. Res. Part C* 18 (4), 568–583.

- Koppa, R.J., 1997. Human factors. In: *Monograph on Traffic Flow Theory*. U.S. Department of Transportation, Federal Highway Administration, Washington, DC, vol. 1997, pp. (3)1–32. <<http://www.tfhrc.gov/its/tft/tft.htm>>.
- Kühne, R., Michalopoulos, P., 1997. Continuum flow models. In: *Monograph on Traffic Flow Theory*. U.S. Department of Transportation, Federal Highway Administration, Washington, DC, vol. 1997, pp. (5)1–51. <<http://www.tfhrc.gov/its/tft/tft.htm>>.
- Kwon, J., McCullough, B., Petty, K., Varaiya, P., 2007. Evaluation of PeMS to improve the congestion monitoring program. California PATH Research Report UCB-ITS-PRR-2007-6, Institute of Transportation Studies, University of California, Berkeley, CA.
- Laval, J., Leclercq, L., 2010. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* 368, 4519–4541.
- Lighthill, M., Whitham, G., 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proc. R. Soc. London Ser. A Math. Phys. Sci.* 229 (1178), 317–345.
- Luspay, T., Kulcsár, B., Varga, I., Zegeye, S.K., De Schutter, B., Verhaegen, M., 2010. On acceleration of traffic flow. In: *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems (ITSC 2010)*, Madeira Island, Portugal, September, 2010, pp. 741–746.
- Montanino, M., Punzo, V., 2013. Making NGSIM data usable for studies on traffic flow theory: a multistep method for vehicle trajectory reconstruction. In: *Proceedings of 92nd TRB Annual Meeting*. Washington, D.C., January 13–17, 2013.
- Payne, H.J., 1971. *Models of Freeway Traffic and Control*. Simulation Council.
- Payne, H.J., 1979. FREFLO: a macroscopic simulation model of freeway traffic. *Transp. Res. Rec.* 722 (1979), 68–75.
- Post, K., Kent, J.H., Tomlin, J., Carruthers, N., 1984. Fuel consumption and emission modelling by power demand and a comparison with other models. *Transp. Res. Part A* 18 (3), 191–213.
- Punzo, V., Borzacchiello, M., Ciuffo, B., 2011. On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data. *Transp. Res. Part C* 19, 1243–1262.
- Richards, P.L., 1956. Shockwaves on the highway. *Oper. Res.* 4 (1), 42–51.
- Siebel, F., Mauser, W., 2006. On the fundamental diagram of traffic flow. *SIAM J. Appl. Math.* 66 (4), 1150–1162.
- Strub, I., Bayen, A., 2006. Weak formulation of boundary conditions for scalar conservation laws: an application to highway traffic modelling. *Int. J. Robust Nonlinear Control*, 16 733–748.
- Tikhonov, A.N., 1963. Regularization of incorrectly posed problems. *Sov. Math. Dokl.* 4, 1624–1627.
- Whitham, G.B., 1974. *Linear and nonlinear waves*. In: *Pure and Applied Mathematics*. Wiley-Interscience, New York.
- Work, D.B., Blandin, S., Tossavainen, O.P., Piccoli, B., Bayen, A.M., 2010. A traffic model for velocity data assimilation. *Appl. Math. Res. Express* 2010 (1), 1–35.
- Yim, Y., Cayford, R., 2001. Investigation of vehicles as probes using global positioning system and cellular phone tracking: field operational test. California PATH Working Paper UCB-ITS-PWP-2001-9, Institute of Transportation Studies, University of California, Berkeley, CA.
- Yuan, Y., Van Lint, J.W.C., Hoogendoorn, S.P., Vrancken, J.L.M., Schreiter, T., 2011. Freeway traffic state estimation using extended Kalman filter for first-order traffic model in Lagrangian coordinates. In: *2011 IEEE International Conference on Networking, Sensing and Control (ICNSC)*. IEEE, pp. 121–126.
- Zhang, H.M., 2002. A non-equilibrium traffic model devoid of gas-like behavior. *Transp. Res. Part B* 36 (3), 275–290.