Use of Speed Measurements for Highway Traffic State Estimation

Case Studies on NGSIM Data and Highway A20, Netherlands

Claudio Roncoli, Nikolaos Bekiaris-Liberis, and Markos Papageorgiou

This paper presents two case studies in which a macroscopic model-based approach for the estimation of traffic conditions, which the authors have recently developed, is employed and tested. The estimation method is developed for a mixed traffic scenario, in which traffic is composed of both ordinary and connected vehicles. Only average speed measurements, which may be obtained from connected vehicle reports, and a minimal number (sufficient to guarantee observability) of spot sensor–based total flow measurements are utilized. In the first case study, NGSIM microscopic data are used to test the capability of estimating traffic conditions on the basis of aggregated information retrieved from moving vehicles and considering various penetration rates of connected vehicles. In the second case study, a longer highway stretch with internal congestion is utilized to test the capability of the proposed estimation scheme to produce appropriate estimates for varying traffic conditions on long stretches. In both cases, the performances are satisfactory and the obtained results demonstrate the effectiveness of the method in both qualitative and quantitative terms.

The estimation of traffic conditions is a task of crucial importance for the development and application of traffic management and control strategies. Essentially, traffic condition estimation consists of estimating the traffic variables of a highway network at sufficient spatial resolution in real time on the basis of a limited amount of information from real-time traffic measurements. In conventional traffic, the necessary measurements are provided by spot sensors (based on a variety of possible technologies), which are placed at specific highway locations. In cases in which the sensor density is sufficiently high (e.g., every 500 m), the collected measurements are usually sufficient for traffic surveillance and control; otherwise, appropriate estimation schemes need to be employed to produce estimates at the required spacial resolution. See, for instance, Munoz et al. (*1*), Wang and Papageorgiou (*2*), Mihaylova et al. (*3*), and Morbidi et al. (*4*), among many other works that address the problem of highway traffic estimation that employs conventional detector data.

In the past two decades there has been a significant and increasing interdisciplinary effort by the automotive industry as well as numerous research institutions around the world to plan, develop, test, and start deploying a variety of vehicle automation and communication systems that are expected to revolutionize the features and capabilities of individual vehicles within the next few decades. A wide description of vehicle automation and communication systems, focusing particularly on their impact on highway traffic efficiency, may be found in Diakaki et al. (*5*). To achieve related traffic flow efficiency improvements on highways, it is of paramount importance to develop novel methods for the modeling, estimation, and control of traffic in the presence of vehicle automation and communication systems (*6–16*).

Some vehicle automation and communication systems are characterized by different types of communication device and therefore provide the possibility of vehicle-to-vehicle or vehicle-to-infrastructure communication; specifically, connected vehicles can report their position, speed, and other relevant information (i.e., they can act as mobile sensors; existing examples include mobile phones and network-connected GPS devices). From a traffic estimation viewpoint, this communication may allow for an improvement in estimation accuracy, as well as a significant reduction in the number of spot sensors needed, which would lead to a reduction of the purchase, installation, and maintenance cost of traffic monitoring.

The exploitation of information retrieved from connected vehicles for traffic estimation purposes is considered in numerous research works. Work et al. designed an ensemble Kalman filter for a Godunov-discretized Lighthill–Whitham–Richards model and estimated the velocity field on highways on the basis of data obtained from GPS devices (*17*); a real-case test was defined on the basis of data retrieved from the Mobile Century experiment (*18*). The work by Treiber et al. was based on data fusion techniques and utilized information retrieved both from fixed and mobile sensors (*19*); two case studies were reported that were based on data from highway A9, near Munich, Germany, and from highway M-42, near Birmingham, United Kingdom. Herrera and Bayen designed a method to incorporate mobile probe measurements into a modified Lighthill–Whitham–Richards model and exploited the so-called Newtonian relaxation (*20*); the efficiency and performance of the estimation algorithm were tested on I-880 in California within the framework of the Mobile Century experiment (*18*). Yuan et al. formulated a traffic model in a Lagrangian coordinate system (in which state variables moved with the traffic stream) that included both Eulerian and Lagrangian measurements (*21*); an extended Kalman filter was then used to obtain appropriate traffic estimates, and the method was tested on data retrieved from highway M-42, near Birmingham. Piccoli et al. addressed the impact of both sampling frequency and penetration rate on the mobile sensing of highway traffic; the study employed a second-order phase transition model and proposed some data fusion schemes to incorporate vehicle trajectory data into the phase transition model (*22*); extensive tests that used NGSIM (Next Generation Simulation) data were performed (*23*). Seo et al. proposed a novel probe vehicle–based estimation method to obtain volume-related

Dynamic Systems and Simulation Laboratory, Department of Production Engineering and Management, Technical University of Crete, Chania 73100, Greece. Corresponding author: C. Roncoli, croncoli@dssl.tuc.gr.

Transportation Research Record: Journal of the Transportation Research Board, No. 2559, Transportation Research Board, Washington, D.C., 2016, pp. 90–100.
DOI: 10.3141/2559-11

variables by assuming that a probe vehicle would measure the spacing to its leading vehicle (*24*); remarkably, the study did not assume knowledge of any fundamental diagram or perform a real experiment employing prototypes of probe vehicles. This method was extended and improved by Seo and Kusakabe, who took into consideration the conservation law (*25*). Bekiaris-Liberis et al. proposed a macroscopic model–based approach for the estimation of the total density and the total flow [i.e., the density or flow of all vehicles (connected and ordinary)] for mixed traffic on the basis of the speed measurements reported by connected vehicles and a few spot flow measurements (*26*); this approach removed the requirement of empirical, thus uncertain, traffic speed modeling, such as the fundamental diagram. The performance of the developed estimation scheme was validated through simulation, with the well-known second-order traffic flow model METANET used as ground truth (*27*).

In this paper, two case studies are presented that use real data retrieved from highways to test and evaluate the behavior of the recently developed estimation scheme (*26*). The two case studies are built on data stemming from different real experiments. The first case, based on NGSIM microscopic data, aims to verify the capability to estimate traffic conditions at high space resolution (i.e., considering segments on the order of 50 m in length) on the basis of aggregate information retrieved from moving vehicles (i.e., the speeds of connected vehicles) and the total flow measured at the entrance and exit of the highway stretch; the proposed estimation scheme is also tested for different penetration rates of connected vehicles. In the second case, a longer highway stretch and a longer time horizon are considered on the basis of information from fixed detectors—namely, speed measurements at every segment (emulating connected vehicle reports) and a small number of flow measurements. This process enables the capability to detect the onset of congestion within the network to be verified; such detection would be one of the most crucial and challenging aspects should this approach be exploited in conjunction with a traffic control strategy. In both cases, the estimation results are analyzed qualitatively and quantitatively through the use of an appropriate metric to compare the performance of the proposed method under different scenarios.

In the following section, a linear parameter-varying model of the density dynamics and a Kalman filter for density estimation are described; then, two case studies to evaluate the performance of the estimation scheme are described. Finally, the paper highlights the main findings and introduces further challenges for future research.

TRAFFIC ESTIMATION ON THE BASIS OF AVERAGE SPEED MEASUREMENTS

Traffic Density Dynamics

The following discrete time equation describes the dynamics of the total densities (ρ_i)—namely, the number of vehicles within segment i divided by the length of the segment (Δ_i)—on subsequent highway segments [e.g., Papageorgiou and Messmer (*27*)]:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{\Delta_i}\left(q_{i-1}(k) - q_i(k) + r_i(k) - s_i(k)\right) \tag{1}$$

where

$i = 1, \ldots, N$ = index of the specific highway segment;
N = number of segments on the highway;
i = variable's value at segment i of the highway;
q_i = total flow at the exit of segment i;
T = time discretization step;
$k = 0, 1, \ldots, k$ = discrete time index, k being the horizon of the estimation; and
r_i and s_i = inflow and outflow of vehicles at on-ramps and off-ramps, respectively, at segment i.

On the basis of the known relationship

$$q_i(k) = \rho_i(k) v_i(k) \tag{2}$$

where v_i is the average speed in segment i, Equation 1 can be rewritten as

$$\rho_i(k+1) = \frac{T}{\Delta_i} v_{i-1}(k)\rho_{i-1}(k) + \left(1 - \frac{T}{\Delta_i} v_i(k)\right)\rho_i(k) + \frac{T}{\Delta_i}\left(r_i(k) - s_i(k)\right) \tag{3}$$

To guarantee that the discrete time relationship described by Equation 2 is sufficiently accurate, the following inequality must be satisfied:

$$\max_{i,k} \frac{\Delta_i}{T} v_i(k) < 1 \tag{4}$$

Under the assumption that the unmeasured on- and off-ramp flows are constant (or slowly varying), the unmeasured ramp flow dynamics may be reflected by a random walk:

$$\theta_i(k+1) = \theta_i(k) + \xi_i^{\theta}(k) \tag{5}$$

where ξ_i^{θ} is zero-mean white Gaussian noise and θ_i is defined in Equation 6:

$$\theta_i = \begin{cases} \dfrac{T}{\Delta_i} r_{n_i} & \text{if } n_i \in L_r \\ \dfrac{T}{\Delta_i} s_{n_i} & \text{if } n_i \in L_r \end{cases} \tag{6}$$

for all $i = 1, \ldots, l_r + l_s$

where $L_r = \{n_1, \ldots, n_{l_r}\}$ and $L_s = \{n_{l_r+1}, \ldots, n_{l_r+l_s}\}$ are the sets of segments, denoted by n_i, that have an on- or off-ramp, respectively, whose flow is not directly measured and l_r and l_s are the number of unmeasured flows at on- and off-ramps, respectively.

It is assumed that at segment i there can be only one on-ramp or only one off-ramp, as is typically the case on a highway. Thus, $L_r \cap L_s = \varnothing$.

Under the assumption that the average speed of vehicles in a segment (v_i) is measured (e.g., from connected vehicle reports) and through the definition of the state shown in Equation 7:

$$x = \left(\rho_1, \ldots, \rho_N, \theta_1, \ldots, \theta_{l_r+l_s}\right)^T \tag{7}$$

the deterministic part of the dynamics of the total density given in Equation 3 and of θ_i given in Equation 5 can be written in the form of a linear parameter-varying system as

$$x = (k+1) = A(k)x(k) + Bu(k) \tag{8}$$

where

$$A(k)=\begin{cases} a_{ij}=\dfrac{T}{\Delta_i}v_{i-1}(k) & \text{if } i-j=1 \text{ and } i\geq 2 \\ a_{ij}=1-\dfrac{T}{\Delta_i}v_i(k) & \text{if } i=j \\ a_{n_i j}=1 & \text{if } n_i\in L_r \text{ and } j=N+i \\ a_{n_i j}=-1 & \text{if } n_i\in L_s \text{ and } j=N+i \\ a_{ij}=1 & \text{if } N<i\leq N_1 \text{ and } j=i \\ a_{ij}=0 & \text{otherwise} \end{cases} \tag{9}$$

$$B=\begin{cases} b_{ij}=\dfrac{T}{\Delta_i} & \text{if } i=1 \text{ and } j=1 \\ b_{m_i j}=\dfrac{T}{\Delta_{m_i}} & \text{if } m_i\notin\bar{L},\ 1\leq m_i\leq N,\ 1\leq i\leq N_2, \text{ and } j=i+1 \\ b_{ij}=0 & \text{otherwise} \end{cases} \tag{10}$$

$$u(k)=\begin{cases} u_i=q_0(k) & \text{if } i=1 \\ u_{i+1}=r_{m_i}-s_{m_i} & \text{if } m_i\notin\bar{L} \end{cases} \tag{11}$$

with $\bar{L}=L_r\cup L_s$, $N_1=N+l_r+l_s$, $N_2=N-l_r-l_s$, $A\in\mathbb{R}^{N_1\times N_1}$, $B\in\mathbb{R}^{N_1\times(N_2+1)}$. The variable q_0, which is assumed to be measured (e.g., by a fixed-flow detector), denotes the total flow of vehicles at the entry of the considered highway stretch and acts as an input to Equation 8, along with any directly measured on- and off-ramp flows, r_i and s_i, $i\notin L_r$, $i\notin L_s$ respectively; v_i, $i=1,\ldots,N$ are viewed as time-varying parameters of Equation 8. Linear parameter-varying systems are a well-studied subclass of linear time-varying systems, whose dynamics vary as a result of the variation of certain parameters [e.g., Sename et al. (*28*)].

For the measured outputs, availability of the density (or, equivalently, the flow) at the mainstream exit of the highway is assumed. If there is exactly one unmeasured ramp within the considered highway stretch, then no additional measurements are necessary for flow observability. However, if there is more than one unmeasured ramp within the stretch, one mainstream flow measurement is needed at any highway segment between every two consecutive unmeasured ramps.

In summary, the measured outputs (y) associated with Equations 8 to 11 are the density (or, equivalently, the flow) at the exit of the considered highway stretch and at a highway segment between every two consecutive ramps whose flows are not measured. Therefore

$$y(k)=Cx(k) \tag{12}$$

where $C\in\mathbb{R}^{(l_r+l_s)\times(N+l_r+l_s)}$ is defined as

$$C=\begin{cases} c_{ij}=1 & \text{for all } i=1,\ldots,l_r+l_s-1 \text{ and some } n_i^*\leq j\leq n_{i+1}^*-1 \\ c_{ij}=1 & \text{if } i=l_r+l_s \text{ and } j=N \\ c_{ij}=0 & \text{otherwise} \end{cases} \tag{13}$$

and $\bar{L}^*=\{n_1^*, n_2^*, \ldots, n_{l_r+l_s}^*\}$ is the set $\bar{L}$ ordered by $<$.

Although it is physically intuitive that the system described in Equations 8 to 13 is observable, when the additional mainstream fixed-flow sensors are placed at some segment between every two consecutive unmeasured ramps, the authors rigorously proved in a previous work that the system was indeed observable for certain cases, such as when a fixed sensor was placed on the mainstream at every segment immediately before an unmeasured ramp (*26*).

The measurement requirements for the proposed estimation algorithm can be summarized as

- The average speed of connected vehicles at a segment of the highway (v_i, $i=1,\ldots,N$), the assumption being that there is no systematic difference between the average speed of the connected vehicles and the average speed of all vehicles in a segment;
- The total flow of vehicles at the entry and exit of the considered highway stretch: q_0 and q_N, respectively; and
- Either the total flow of vehicles at a ramp (e.g., r_i or s_i) or an additional mainstream flow measurement (e.g., q_j) at any highway segment between two consecutive unmeasured ramps.

Kalman Filter

A Kalman filter is employed for the estimation of the total density of vehicles in each segment of a highway (Figure 1). Defining the estimated state vector as

$$\hat{x}=\left(\hat{\rho}_1,\ldots,\hat{\rho}_N,\hat{\theta}_1,\ldots,\hat{\theta}_{l_r+l_s}\right)^T \tag{14}$$

the Kalman filter equations [e.g., Anderson and Moore (*29*)] are

$$\hat{x}(k+1)=A(k)\hat{x}(k)+Bu(k)+A(k)G(k)\left(z(k)-C\hat{x}(k)\right) \tag{15}$$

$$G(k)=P(k)C^T\left(CP(k)C^T+R\right)^{-1} \tag{16}$$

$$P(k+1)=A(k)\left(I-G(k)C\right)P(k)A(k)^T+Q \tag{17}$$

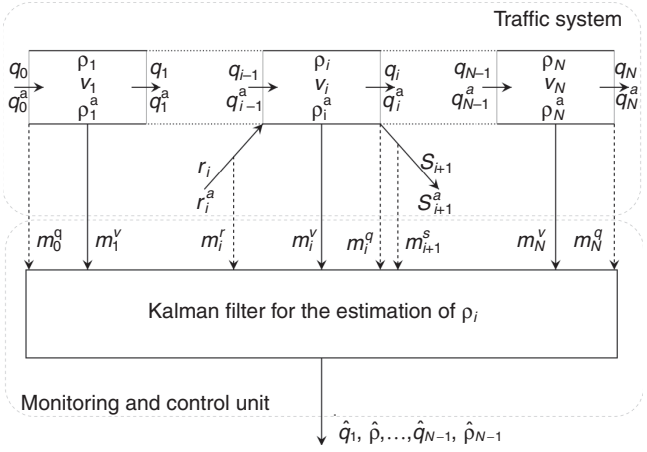


FIGURE 1 Traffic system and proposed Kalman filter, which operates on speed measurements from connected vehicles (solid lines) and flow measurements from fixed sensors (dashed lines) (m_i^w = measurement of quantity w at segment i; w_i^a = quantity w of connected vehicles at segment i; m_i^w may differ from actual quantity w, for example because of measurement noise).

where

G = Kalman gain,
P = predicted estimated covariance,
z = noisy version of measurement y,
$R = R^T > 0$ and $Q = Q^T > 0$ = tuning parameters, and
I = identity matrix of size $N + I\_r + I\_s$.

In the ideal case in which there is additive, zero-mean Gaussian white noise in Equations 8 and 12, R and Q, respectively, represent the (ideally known) covariance matrices of the measurement and process noises. (In particular, the $Q_{N+i,N+i}$ elements of Q represent the filter's anticipation for the covariance of ξ_i^{θ}.) Since the system equations here are relatively complex, some tuning of R and Q may be needed for the best estimation results. The initial conditions of the filter described by Equations 15 to 17 are chosen as

$$\hat{x}(k_0) = \mu \tag{18}$$

$$P(k_0) = H \tag{19}$$

where μ and $H = H^T > 0$, in the ideal case in which $x(k_0)$ is a Gaussian random variable, represent the mean and auto covariance matrix of $x(k_0)$, respectively. The Kalman filter delivers estimates of the segment total densities ($\hat{\rho}_i$), as indicated at the output of the Kalman filter in Figure 1.

EXPERIMENTAL TESTING OF THE PROPOSED ESTIMATION METHOD

Case Study 1. NGSIM I-80 Data

Network and Data Description

To evaluate and illustrate the effectiveness of the proposed estimation scheme, microscopic traffic data collected within the NGSIM program were utilized (*23*). However, since these data incorporate nonnegligible errors in the position of individual vehicles (*30*), some correction methods have been proposed to improve their reliability (*31, 22*). This work utilized the data processed by Montanino and Punzo (*31, 32*), which included the trajectories of all vehicles traveling along a stretch in the northbound direction of I-80 in Emeryville, California, recorded from 4:00 to 4:15 p.m. on April 13, 2005. The highway was composed of six lanes; however, Lane 1 was a high-occupancy vehicle lane, on which access was restricted to vehicles with a driver and one or more passengers. The lane's traffic characteristics were therefore structurally different than those of the other lanes; therefore, this lane was excluded from the estimation. The considered stretch (shown in Figure 2) was 400 m long, with an on-ramp entering the mainstream; the merge nose was located 175 m after the network origin. As can be seen from Figure 3, massive congestion was present within the stretch; congestion waves came from downstream and crossed the entire stretch.

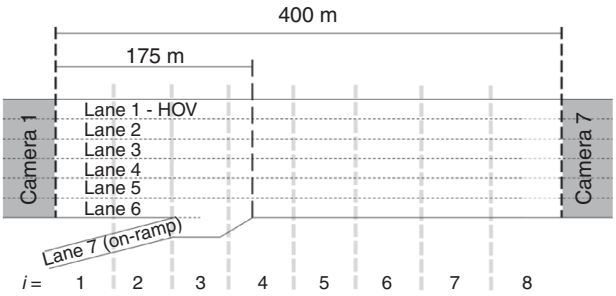


FIGURE 2 Stretch of I-80 in Emeryville used in Case Study 1.

To perform macroscopic evaluations, the stretch was divided into $N = 8$ homogeneous segments of 50 m in length (see Figure 2); the on-ramp was placed within Segment 4, and a discrete time step ($T = 5$ s) was used. Even though this configuration slightly violated the condition of Equation 4 for some time intervals and some segments, the obtained results were very similar to those obtained when lower values that fully satisfied the condition of Equation 4 were used for the time step T.

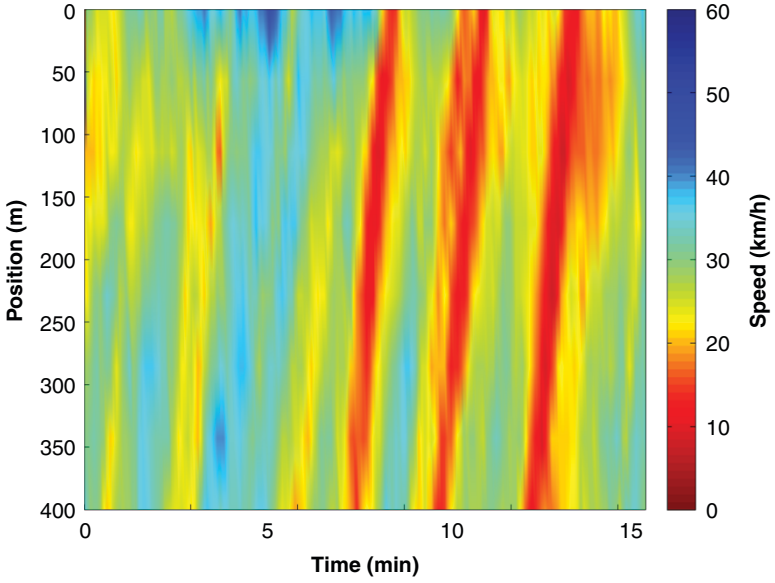


FIGURE 3 Aggregated vehicle speeds extracted from NGSIM data.

All necessary information to perform estimation was extracted from the available trajectory data. In particular, the segment speed $v_i(k)$ was computed as the arithmetic average of the speeds of the connected vehicles present within the segment at the measurement time instant kT. Total flow measurements were computed via two virtual spot detectors: one placed at the network entrance (beginning of segment $i = 1$) to provide q_0, and another placed at the network exit (end of segment $i = 8$) to provide q_N. The corresponding flow values were computed by counting the number of vehicles (irrespective of being connected or not) that crossed the virtual detector locations during the period.

The estimation scheme was configured to estimate the mainstream densities and the total on-ramp flow. In the first scenario, it was assumed that all vehicles were connected; thus, the segment speed was computed on the basis of information from all the vehicles present. In the second scenario, variable percentages of connected vehicles were assumed to be present on the highway; thus, the estimation algorithm used the less accurate measurement of the segment speeds. In the latter case, vehicles entering the network were randomly marked as connected according to the assumed penetration rate and on the basis of a uniform distribution.

The ground truth, used to evaluate the estimation results, was represented by the total density in each segment and the on-ramp flow. The segment densities $[\rho_i(k)]$ were computed by counting the number of vehicles present within segment i at time instant kT, divided by the segment length (0.05 km); the on-ramp flow $r_4(k)$ was computed similarly to the computation of flow by a virtual detector, that is, by counting the number of vehicles leaving the on-ramp (Lane 7 in Figure 3) and entering the mainstream during time interval $(k, k + 1]$.

Performance Evaluation Under 100% Penetration Rate

In the first scenario, it was assumed that all vehicles were connected and thus capable of communicating their speed and position to the central authority; consequently, the filter was fed with accurate measurements of the average segment speeds. The parameters used for the Kalman filter were

$$Q_{i,i} = 1, 1 \le i \le N; Q_{i,i} = 0.01, N < i \le N + l_r + l_s;$$

$$Q_{i,j} = 0, \forall\, i \ne j; R = 10\, I_{(l_r+l_s)\times(l_r+l_s)}$$

$$\mu = (40, \ldots, 40); H = I_{(N+l_r+l_s)\times(N+l_r+l_s)}$$

Under a 100% penetration rate, the filter would not be really needed, because the total densities and flows could be directly calculated on the basis of the vehicle position report. Nevertheless, the estimation scheme could also be applied to this case to enable comparison with lower penetration rate results.

The following performance index, formulated as the coefficient of variation (CV in Equation 17) of the root mean square error of the estimated density $\hat{\rho}$ with respect to the ground truth density ρ, is proposed for the numerical evaluation of the estimation method:

$$CV_\rho = \frac{\sqrt{\frac{1}{KN}\sum_{k=1}^{K}\sum_{i=1}^{N}\left[\hat{\rho}_i(k) - \rho_i(k)\right]^2}}{\frac{1}{KN}\sum_{k=1}^{K}\sum_{i=1}^{N}\rho_i(k)} \tag{20}$$

The obtained results are characterized by $CV_\rho = 14.9\%$, which implies a very accurate estimation; this accuracy is also visible in Figure 4, which shows that the produced estimates capture the dynamic variations of traffic conditions with high accuracy.

The estimation of the ramp flow suffers from some delay that is caused by the distance of the downstream detector from the on-ramp location; in fact, the dynamics of flow entering from the on-ramp can be captured only after the on-ramp vehicles reach the measurement location (i.e., the end of the stretch). These discrepancies could likely be substantially reduced by placing an additional flow detector immediately downstream of the merging segment.

Performance Evaluation Under Various Penetration Rates

Within corresponding experiments, mixed traffic (connected and ordinary vehicles) was considered with different penetration rates of connected vehicles. Essentially, this process evaluated the robustness of the approach to less-accurate measurements of the segment speeds. In this case, because of the limited number of vehicles transmitting information, two main issues could arise: first, the measurement retrieved from a limited number of vehicles may not be fully representative of the real average speed of a segment; second, there may be no connected vehicles within some segments when the measurements are retrieved. Potential reasons for the former are related to microscopic and behavioral phenomena; for example, consider a case in which there is only one connected vehicle in a segment and the vehicle's speed is retrieved after the driver is forced to strongly decelerate because of an aggressive lane-changing maneuver by some other vehicle or a case in which the connected vehicle is driven by a particularly fast (or slow) driver. To compensate for potential large errors in the reported segment speeds in such situations, a moving average of the reported segment speeds is employed to feed the filter, rather than the use of instantaneous segment speeds. In particular, the last three measurements are averaged (thus, considering a time window of 15 s), ignoring measurements at time steps when no speed is reported (because no connected vehicles are present). The choice of the time window is based on some numerical experiments. When choosing this value, one should account for a trade-off between smoothness (a large time window) and delay (short time windows); nevertheless, the conducted experiments indicated that the related performance sensitivity of the results was rather low.

The parameters used for the Kalman filter were kept the same as in the previous case. From the performance index comparison displayed in Figure 5, it appears that the use of a moving average for the reported segment speeds was indeed advantageous over the use of instantaneous speeds for low penetration rates and that the performance of the estimation algorithm slowly degraded with a decrease in the penetration rate. Thus, the estimation scheme is seen to produce good results even if the penetration rate of connected vehicles is very low (e.g., 2%). For the sake of brevity, only the results related to one of these tests—namely, with a 5% penetration rate of connected vehicles and when feeding the filter with the moving average of the reported speeds during the last 15 s—are illustrated in Figure 6; despite the obviously less accurate estimation compared with Figure 4, it can be seen that the estimator is capable of following the real density pattern and the on-ramp flow dynamics (with some delay).

Tests were also performed to evaluate the magnitude of the error related to the incorrect measurement of average segment speeds for

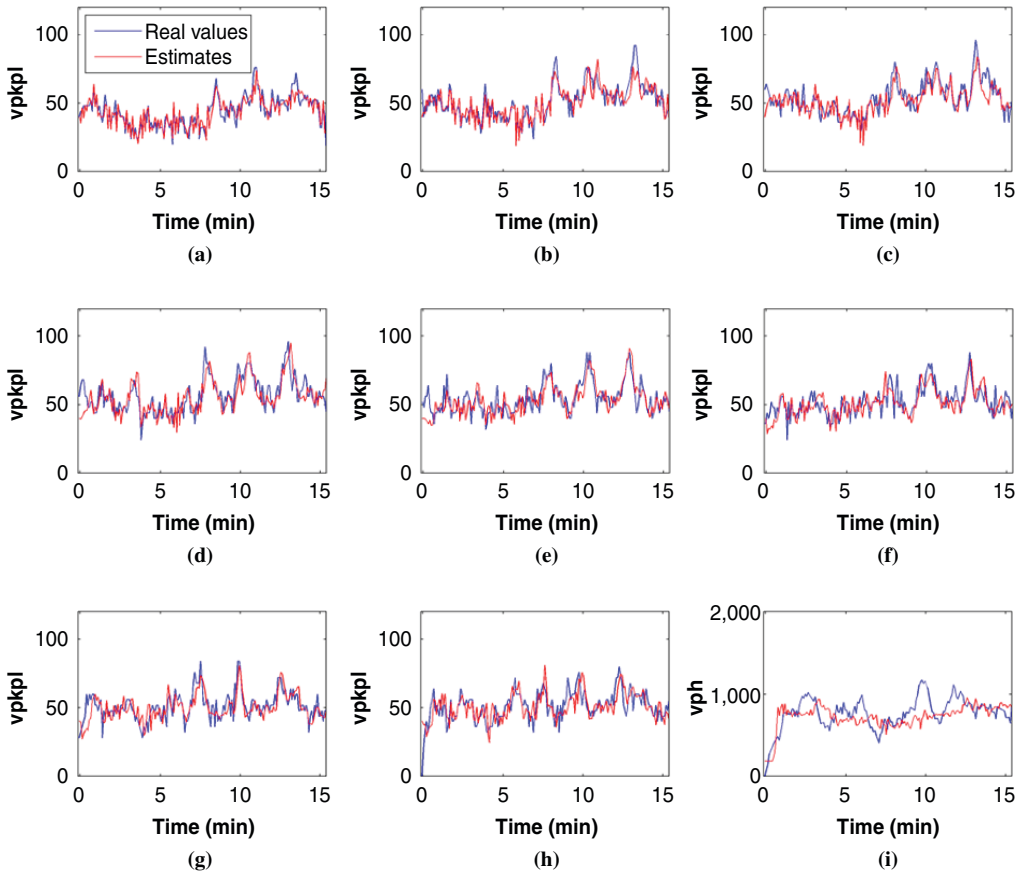


FIGURE 4 Trajectories of real and estimated densities for NGSIM case study, under assumption that all vehicles are connected with perfect knowledge of segment speeds: (*a*) Segment 1, (*b*) Segment 2, (*c*) Segment 3, (*d*) Segment 4, (*e*) Segment 5, (*f*) Segment 6, (*g*) Segment 7, (*h*) Segment 8, and (*i*) on-ramp flow (vpkpl = vehicles per kilometer per lane; vph = vehicles per hour).

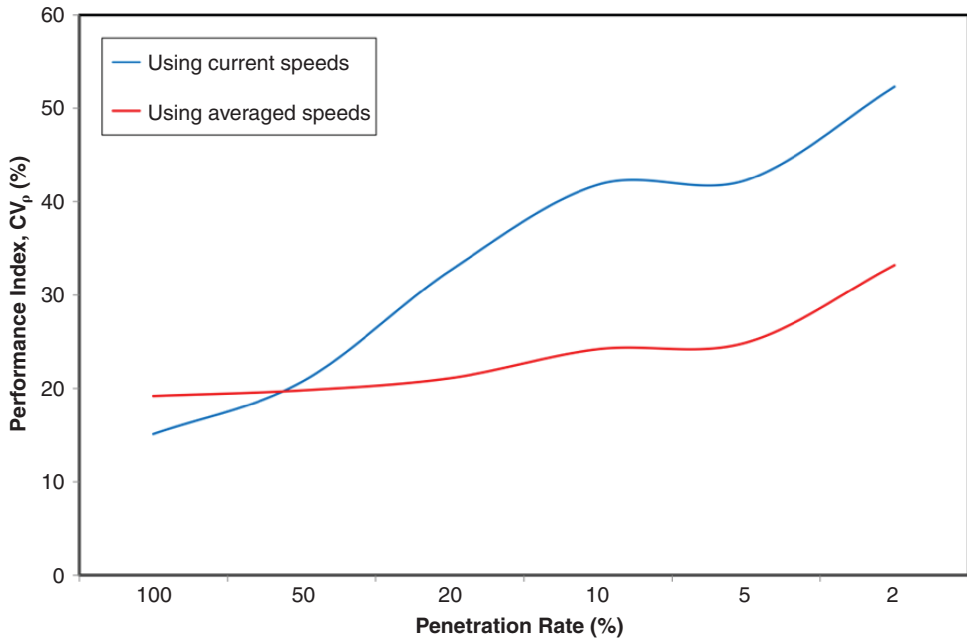


FIGURE 5 Comparison of performance indexes of relative error for NGSIM data, considering different penetration rates of connected vehicles and using current speeds and a moving average of reported speeds during preceding 15 s.

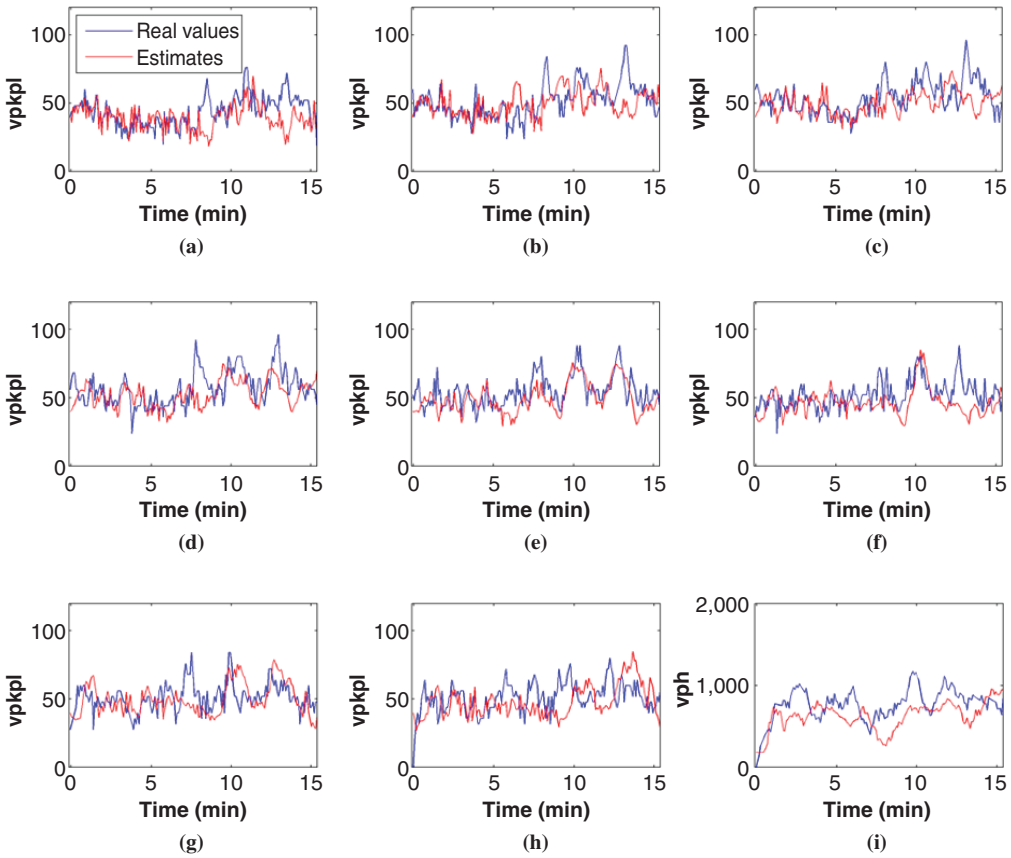


FIGURE 6 Trajectories of real and estimated densities for NGSIM case study, under assumption of 5% penetration rate of connected vehicles: (*a*) Segment 1, (*b*) Segment 2, (*c*) Segment 3, (*d*) Segment 4, (*e*) Segment 5, (*f*) Segment 6, (*g*) Segment 7, (*h*) Segment 8, and (*i*) on-ramp flow (filter fed with a moving average of reported speeds during past 15 s).

cases with a low penetration rate of connected vehicles. In particular, 10 simulations were run that considered different sets of vehicles being connected (for a fixed penetration rate), then computing the mean covariance (w) of the expected value ($E[\cdot]$) of the conservation equation error attributable to a discrepancy between the segment speed computed on the basis of information from connected vehicles [$\hat{v}_i(k)$] and the real average segment speed [$\overline{v}_i(k)$], according to

$$w = \frac{1}{KN}\sum_{i=1}^{N}\sum_{k=1}^{K}\frac{T^2}{\Delta_i^2}E\left\{\rho_i^2(k)\left[\hat{v}_i(k)-\overline{v}_i(k)\right]^2\right\} \tag{21}$$

The resulting value for a penetration rate of connected vehicles equal to 20% is $w = 8.1$ vehicles^2/km^2; for a penetration rate of 5%, the resulting mean covariance is $w = 19.2$ vehicles^2/km^2. These values suggest that better performance may be achieved through the use of higher values of Q for lower penetration rates. Nevertheless, experimental tests (not reported here) indicate that for low penetration rates, the performance of the proposed estimation scheme does not vary significantly for different values of Q.

Unfortunately, NGSIM data were collected for a short highway stretch within a peak period, characterized by congested traffic conditions only; therefore, the capability of the proposed estimation scheme to capture the flow breakdown occurrence cannot be tested with these data. To extend the evaluation range of the proposed estimation scheme, another case study was performed, and it is presented in the next section.

Case Study 2. Highway A20, Netherlands

Network and Data Description

The second case study was based on real data—taken from Schakel and van Arem (*33*)—obtained from detectors on a stretch of the highway A20 between Rotterdam and Gouda in the Netherlands. The topological characteristics of this network—which incorporated a nontrivial combination of lane drops, on-ramps, and off-ramps—the congestion pattern, and the relatively closely spaced detectors made it a challenging test bed for the estimation scheme proposed in the previous section.

The considered stretch was about 11 km in length and included four on-ramps and three off-ramps (however, the gas station on- and off-ramp were ignored since they were characterized by extremely low flows). The stretch included a lane drop at 5,125 m (see Figure 7). There were 32 detectors placed rather homogeneously at a distance of 300 m on average, as illustrated in Figure 7. In addition, a flow detector was located at the off-ramp marked "Moordrecht."

Data from the morning rush hour of Monday, June 8, 2009, were employed; strong congestion occurred around 6:20 a.m. because of the increased flow entering from the on-ramp marked "Nieuwerkerk a/d IJssel"; consequently, the congestion spilled back and worsened at the lane drop area and stretched to the gas station area. The congestion disappeared after 7:40 a.m. because of reduced demand. A three-dimensional plot that illustrates the traffic conditions from 5 to 9 a.m. is shown in Figure 8*a*. This congestion pattern allowed

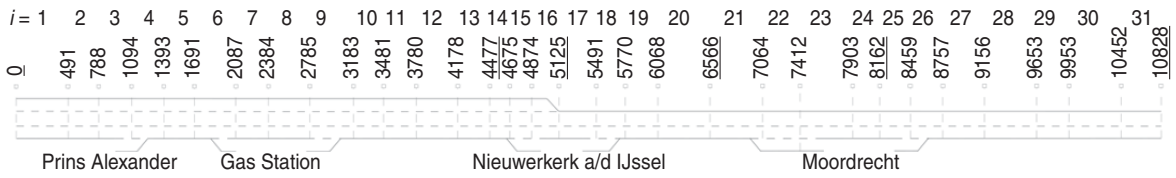
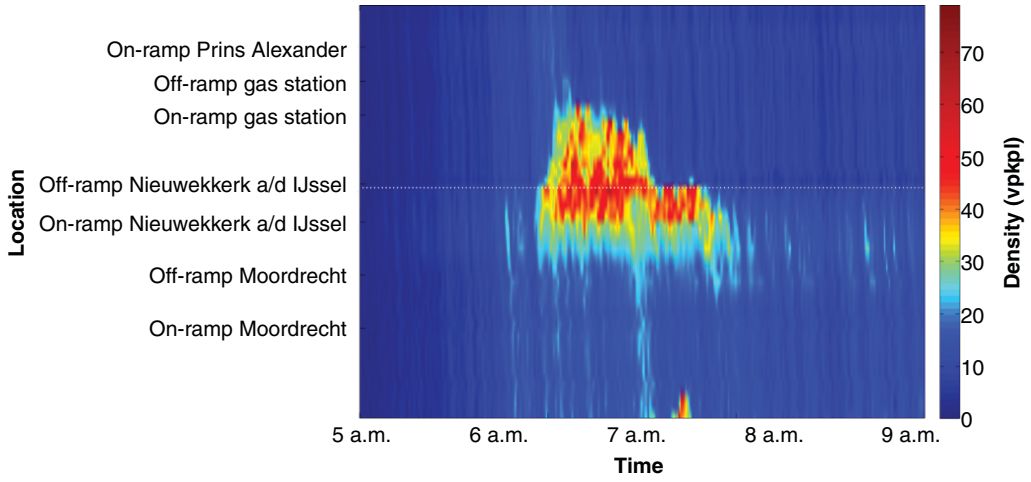
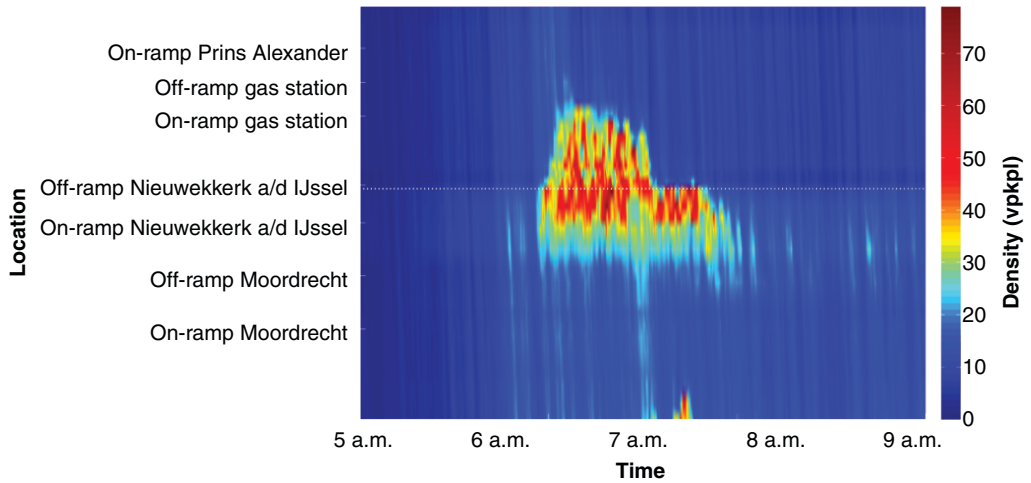


FIGURE 7 Considered stretch of A20 from Rotterdam to Gouda. [Detector positions indicated by distance (m) from network entrance. Detectors used to obtain flow measurements within case study are underlined.]



(a)



(b)

FIGURE 8 Space–time evolution of (*a*) real and (*b*) estimated densities for Case Study 2 (congestion originating at the Nieuwerkerk a/d IJssel merge area marked as dotted line).

the proposed estimator to be tested and evaluated under varying traffic conditions, which included the formation and dissipation of stretch-internal congestion, which was not visible at the stretch boundaries.

The network was space discretized with the definition of $N = 31$ segments, and each segment was delimited by a pair of detectors. According to this space discretizaton, on-ramps were placed within Segments 3, 17, and 25, and off-ramps were placed within Segments 14 and 21 (see Figure 7).

The proposed estimation algorithm was fed with the following information: (*a*) flow measurements retrieved from a limited number of detectors, according to the configuration shown in Figure 7, sufficient to guarantee observability as explained in the previous section [for details, see Bekiaris-Liberis et al. (*26*)], and (*b*) speed measurements retrieved from all detectors, under the assumption that similar information may be obtained from connected vehicle reports. Specifically, the detector located at the highway entrance (0 m) was used to obtain the input of the system q_0; all ramp flows were assumed to be unknown; therefore, four additional detectors were utilized (one between every pair of unmeasured ramps) at 4,477 m (segment $i = 13$), 5,125 m (segment $i = 16$), 6,566 m (segment $i = 20$), and 8,162 m (segment $i = 24$). Finally, one measurement was taken at the network exit (10,828 m) and employed as q_N (see Figure 7).

The ground truth was represented by the densities computed on the basis of the measurements from each detector as $\rho = q/v$, where q and v are the measured flow and speed, respectively, and by the only ramp flow measurement available (the flow at off-ramp Moordrecht). The use of the same measurements to both feed the estimator and construct the ground truth implied that the estimator was not subject to any measurement error. However, to assess the performance of the estimation scheme in more realistic scenarios, two additional cases were considered. In the first, a zero-mean Gaussian white additive noise, characterized by a standard deviation of 300 vehicles per hour (vph), was added to the flow measurements that were utilized by the estimator; in the second, a zero-mean Gaussian white noise, with a standard deviation of 5 km/h, was added to the speed measurements that were employed by the filter, in addition to the flow measurement noise. Even though additive Gaussian noise only may not be always realistic, being potentially a limitation for the effectiveness of a conventional Kalman filter–based estimation approach, a series of works presented in the literature have shown that the Kalman filter may be robust to nonadditive or non-Gaussian noise or have developed robust versions of the Kalman filter for such cases [e.g., Petersen and Savkin (*34*)].

Performance Evaluation

The parameters used for the Kalman filter were

$$Q_{i,i} = 1, 0 \le i \le N; Q_{i,i} = 0.01, N < i \le N + l_r + l_s;$$

$$Q_{i,j} = 0, \forall i \ne j; R = 100\, I_{(l_r+l_s)\times(l_r+l_s)}$$

$$\mu = (4, \ldots, 4); H = I_{(N+l_r+l_s)\times(N+l_r+l_s)}$$

The performance index in Equation 20 was used for the numerical evaluation of the estimation method. In all the tested cases, the estimates were very accurate. In the first case, in which noise was not added to the measurements utilized by the estimator, a performance index value of $CV_\rho = 14.6\%$ was obtained. When the flow measurements used by the filter were subject to noise, a value of $CV_\rho = 15.0\%$ was obtained, which was only a slight degradation from the noise-free case, since the proposed estimation scheme was capable of efficiently handling the effect of additive noise. However, when noise was also added to the speed measurements, a more significant deterioration of the performance index ($CV_\rho = 16.1\%$) and noisier estimates were obtained; this result is attributable to noise, in this case, not being purely additive but also multiplicative within the dynamic equations of the filter. In the three-dimensional plot shown in Figure 8*b*, which illustrates the estimated densities when noise is added to both flow and speed measurements, it can be seen that the estimator reliably reconstructed the congestion pattern within the network in both time and space. Furthermore, Figure 9 depicts the measured and the estimated density trajectories around the location at which traffic congestion began, as well as the flow estimation at off-ramp Moordrecht (the only one for which flow measurements were actually available). It can be seen that the estimation scheme captured the onset of congestion with accurate timing and at the correct location, as well as produced an accurate estimate of the off-ramp flow.

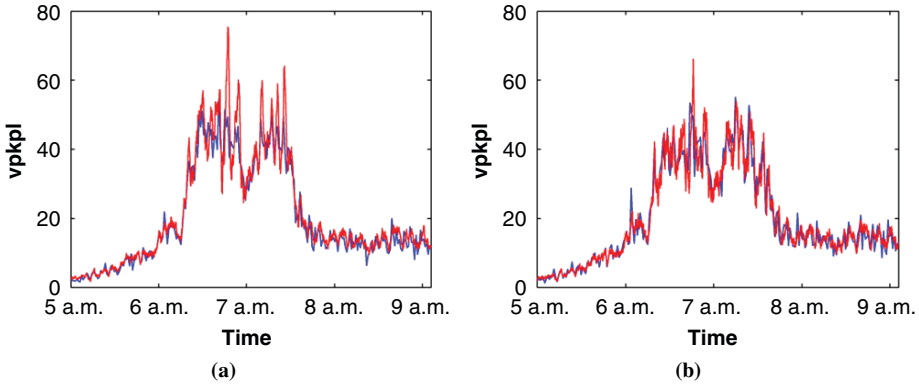


FIGURE 9 Real and estimated density trajectories at location where congestion starts and the flow at off-ramp Moordrecht, for Case Study 2, when additive noise affects both flow and speed measurements: (*a*) detector located at 5,125 m and (*b*) detector located at 5,491 m.

(continued)

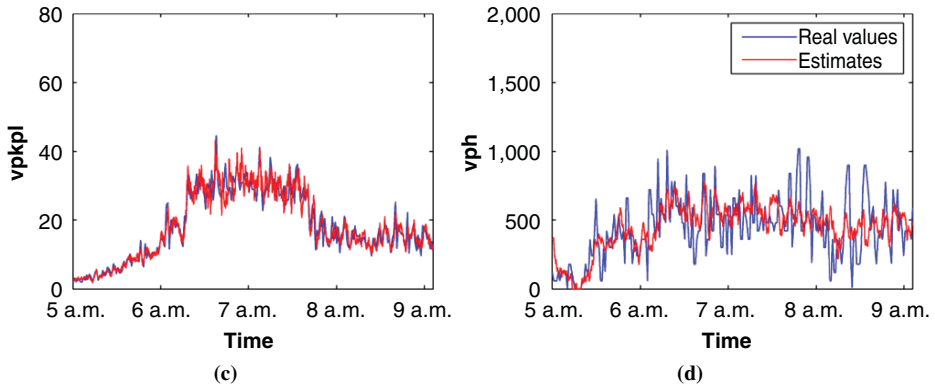


FIGURE 9 *(continued)* Real and estimated density trajectories at location where congestion starts and the flow at off-ramp Moordrecht, for Case Study 2, when additive noise affects both flow and speed measurements: (*c*) detector located at 5,770 m and (*d*) off-ramp Moordrecht.

CONCLUSIONS

The macroscopic model–based traffic estimation scheme proposed in Bekiaris-Liberis et al. (*26*) has been tested on two case studies that used data from different real experiments. In both cases, the results demonstrated the effectiveness of the method in both qualitative and quantitative terms.

Ongoing work involves microscopic simulations to assess the effectiveness of the approach in cases in which connected vehicles are characterized by significantly different behavior (e.g., when they are also automated).

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme. The authors thank Vincenzo Punzo and Marcello Montanino of the University of Naples, Italy, for their support in providing the reconstructed NGSIM data used in the first case study and Bart van Arem and Wouter Schakel of the Delft University of Technology, Netherlands, for the data and information related to the network used in the second case study.

REFERENCES

1. Munoz, L., X. Sun, R. Horowitz, and L. Alvarez. Traffic Density Estimation with the Cell Transmission Model. In *Proceedings of the American Control Conference,* American Automatic Control Council, 2003, pp. 3750–3755.
2. Wang, Y., and M. Papageorgiou. Real-Time Freeway Traffic State Estimation Based on Extended Kalman Filter: A General Approach. *Transportation Research Part B: Methodological,* Vol. 39, No. 2, 2005, pp. 141–167.
3. Mihaylova, L., R. Boel, and A. Hegyi. Freeway Traffic Estimation Within Particle Filtering Framework. *Automatica,* Vol. 43, No. 2, 2007, pp. 290–300.
4. Morbidi, F., L. Ojeda, C. deWit, and I. Bellicot. A New Robust Approach for Highway Traffic Density Estimation. In *Proceedings of the European Control Conference,* European Control Association, 2014, pp. 2575–2580.
5. Diakaki, C., M. Papageorgiou, I. Papamichail, and I. Nikolos. Overview and Analysis of Vehicle Automation and Communication Systems from a Motorway Traffic Management Perspective. *Transportation Research Part A: Policy and Practice,* Vol. 75, 2015, pp. 147–165.
6. Varaiya, P. Smart Cars on Smart Roads: Problems of Control. *IEEE Transactions on Automatic Control,* Vol. 38, No. 2, 1993, pp. 195–207.
7. Rao, B., and P. Varaiya. Roadside Intelligence for Flow Control in an Intelligent Vehicle and Highway System. *Transportation Research Part C: Emerging Technologies,* Vol. 2, No. 1, 1994, pp. 49–72.
8. Rajamani, R., and S. Shladover. An Experimental Comparative Study of Autonomous and Co-Operative Vehicle–Follower Control Systems. *Transportation Research Part C: Emerging Technologies,* Vol. 9, No. 1, 2001, pp. 15–31.
9. Bose, A., and P. Ioannou. Mixed Manual/Semi-Automated Traffic: A Macroscopic Analysis. *Transportation Research Part C: Emerging Technologies,* Vol. 11, No. 6, 2003, pp. 439–462.
10. van Arem, B., C. van Driel, and R. Visser. The Impact of Cooperative Adaptive Cruise Control on Traffic-Flow Characteristics. *IEEE Transactions on Intelligent Transportation Systems,* Vol. 7, No. 4, 2006, pp. 429–436.
11. Kesting, A., M. Treiber, M. Schönhof, and D. Helbing. Adaptive Cruise Control Design for Active Congestion Avoidance. *Transportation Research Part C: Emerging Technologies,* Vol. 16, No. 6, 2008, pp. 668–683.
12. Shladover, S. E., D. Su, and X.-Y. Lu. Impacts of Cooperative Adaptive Cruise Control on Freeway Traffic Flow. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2324,* Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 63–70.
13. Ge, J. I., and G. Orosz. Dynamics of Connected Vehicle Systems with Delayed Acceleration Feedback. *Transportation Research Part C: Emerging Technologies,* Vol. 46, 2014, pp. 46–64.
14. Wang, M., W. Daamen, S. P. Hoogendoorn, and B. van Arem. Rolling Horizon Control Framework for Driver Assistance Systems. Part II: Cooperative Sensing and Cooperative Control. *Transportation Research Part C: Emerging Technologies,* Vol. 40, 2014, pp. 290–311.
15. Roncoli, C., M. Papageorgiou, and I. Papamichail. Traffic Flow Optimisation in Presence of Vehicle Automation and Communication Systems—Part II: Optimal Control for Multi-Lane Motorways. *Transportation Research Part C: Emerging Technologies,* Vol. 57, 2015, pp. 260–275.
16. Roncoli, C., I. Papamichail, and M. Papageorgiou. Model Predictive Control for Multi-Lane Motorways in Presence of VACS. In *Proceedings of the IEEE International Conference on Intelligent Transportation Systems,* IEEE, New York, 2014, pp. 501–507.
17. Work, D., O.-P. Tossavainen, S. Blandin, A. Bayen, T. Iwuchukwu, and K. Tracton. An Ensemble Kalman Filtering Approach to Highway Traffic Estimation Using GPS Enabled Mobile Devices. In *Proceedings of the IEEE Conference on Decision and Control,* IEEE, New York, 2008, pp. 5062–5068.
18. Herrera, J. C., D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen. Evaluation of Traffic Data Obtained Via GPS-Enabled Mobile Phones: The Mobile Century Field Experiment. *Transportation Research Part C: Emerging Technologies,* Vol. 18, No. 4, 2010, pp. 568–583.

19. Treiber, M., A. Kesting, and R.E. Wilson. Reconstructing the Traffic State by Fusion of Heterogeneous Data. *Computer-Aided Civil and Infrastructure Engineering,* Vol. 26, No. 6, 2011, pp. 408–419.
20. Herrera, J.C., and A.M. Bayen. Incorporation of Lagrangian Measurements in Freeway Traffic State Estimation. *Transportation Research Part B: Methodological,* Vol. 44, No. 4, 2010, pp. 460–481.
21. Yuan, Y., J. van Lint, R. Wilson, F. van Wageningen-Kessels, and S. Hoogendoorn. Real-time Lagrangian Traffic State Estimator for Freeways. *IEEE Transactions on Intelligent Transportation Systems,* Vol. 13, No. 1, 2012, pp. 59–70.
22. Piccoli, B., K. Han, T.L. Friesz, T. Yao, and J. Tang. Second-Order Models and Traffic Data from Mobile Sensors. *Transportation Research Part C: Emerging Technologies,* Vol. 52, 2015, pp. 32–56.
23. FHWA, U.S. Department of Transportation. Next Generation Simulation (NGSIM). 2005. www.ngsim-community.org.
24. Seo, T., T. Kusakabe, and Y. Asakura. Estimation of Flow and Density Using Probe Vehicles with Spacing Measurement Equipment. *Transportation Research Part C: Emerging Technologies,* Vol. 53, 2015, pp. 134–150.
25. Seo, T., and T. Kusakabe. Probe Vehicle-Based Traffic State Estimation Method with Spacing Information and Conservation Law. *Transportation Research Part C: Emerging Technologies,* Vol. 59, 2015, pp. 391–403.
26. Bekiaris-Liberis, N., C. Roncoli, and M. Papageorgiou. Highway Traffic State Estimation with Mixed Connected and Conventional Vehicles Using Speed Measurements. In *Proceedings of the IEEE International Conference on Intelligent Transportation Systems,* IEEE, New York, 2015, pp. 2806–2811.
27. Papageorgiou, M., and A. Messmer. METANET: A Macroscopic Simulation Program for Motorway Networks. *Traffic Engineering and Control,* Vol. 31, No. 9, 1990, pp. 466–470.
28. Sename, O., P. Gaspar, and J. Bokor (eds.). *Robust Control and Linear Parameter Varying Approaches,* Vol. 437 of *Lecture Notes in Control and Information Sciences.* Springer, Berlin, 2013.
29. Anderson, B., and J. Moore. *Optimal Filtering.* Prentice Hall, Englewood Cliffs, N.J., 1979.
30. Punzo, V., M.T. Borzacchiello, and B. Ciuffo. On the Assessment of Vehicle Trajectory Data Accuracy and Application to the Next Generation Simulation (NGSIM) Program Data. *Transportation Research Part C: Emerging Technologies,* Vol. 19, No. 6, 2011, pp. 1243–1262.
31. Montanino, M., and V. Punzo. Making NGSIM Data Usable for Studies on Traffic Flow Theory: Multistep Method for Vehicle Trajectory Reconstruction. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2390,* Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 99–111.
32. Montanino, M., and V. Punzo. Reconstructed NGSIM I80-1: Cost Action TU0903—MULTITUDE. 2013. www.multitude-project.eu/exchange/101.html.
33. Schakel, W.J., and B. van Arem. Improving Traffic Flow Efficiency by In-Car Advice on Lane, Speed, and Headway. *IEEE Transactions on Intelligent Transportation Systems,* Vol. 15, No. 4, 2014, pp. 1597–1606.
34. Petersen, I.R., and A.V. Savkin. *Robust Kalman Filtering for Signals and Systems with Large Uncertainties.* Birkhäuser, Boston, Mass., 1999.

The Standing Committee on Vehicle–Highway Automation peer-reviewed this paper.