



BAYESIAN ANALYSIS OF TRAFFIC FLOW ON INTERSTATE I-55: THE LWR MODEL

Author(s): Nicholas Polson and Vadim Sokolov

Source: *The Annals of Applied Statistics*, December 2015, Vol. 9, No. 4 (December 2015), pp. 1864-1888

Published by: Institute of Mathematical Statistics

Stable URL: <https://www.jstor.org/stable/43826447>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve and extend access to *The Annals of Applied Statistics*

BAYESIAN ANALYSIS OF TRAFFIC FLOW ON INTERSTATE I-55: THE LWR MODEL

BY NICHOLAS POLSON AND VADIM SOKOLOV

University of Chicago and Argonne National Laboratory

Transportation departments take actions to manage traffic flow and reduce travel times based on estimated current and projected traffic conditions. Travel time estimates and forecasts require information on traffic density which are combined with a model to project traffic flow such as the Lighthill–Whitham–Richards (LWR) model. We develop a particle filtering and learning algorithm to estimate the current traffic density state and the LWR parameters. These inputs are related to the so-called fundamental diagram, which describes the relationship between traffic flow and density. We build on existing methodology by allowing real-time updating of the posterior uncertainty for the critical density and capacity parameters. Our methodology is applied to traffic flow data from interstate highway I-55 in Chicago. We provide a real-time data analysis of how to learn the drop in capacity as a result of a major traffic accident. Our algorithm allows us to accurately assess the uncertainty of the current traffic state at shock waves, where the uncertainty is a mixture distribution. We show that Bayesian learning can correct the estimation bias that is present in the model with fixed parameters.

1. Introduction. Effectively managing traffic flow to reduce congestion can improve communities by reducing travel times, reducing pollution and improving economic efficiency. Transportation departments use information on current and projected travel times to adjust ramp metering and traffic lights; travelers use projected travel times to make travel plans and to adjust departure times, transportation mode and route. Estimated travel times are developed using sophisticated models of traffic flow that begin with observations on speed and density and develop estimates of road capacity based on estimates of current density and flow.

In their seminal paper, Lighthill and Whitham (1955) describe the theory of kinematic wave motion which they apply to modeling highway traffic flow. Richards (1956) independently proposed a similar application. The key assumption is a relationship between traffic flow and density. A model is calibrated using the characteristics of road segments, such as the number of lanes, free-flow speed and road type. These characteristics themselves do not explain all the variation in model parameters and estimates need to be assessed using observations on current speed, density and lane configurations at sparse points throughout the

Received February 2015; revised September 2015.

Key words and phrases. Traffic flow, intelligent transportation system, LWR model, particle filtering, Bayesian posterior, traffic prediction.

network. Much of the recent improvement in travel time estimation and forecasting has come from improving the estimates of network characteristics [Dervisoglu et al. (2009), Muralidharan and Horowitz (2009)].

Usually, underlying traffic data is sparse. We observe specific points in a traffic network using fixed loop-sensors or at random points via GPS-equipped probe vehicles. Underlying road capacity might vary [Brilon, Geistefeldt and Regler (2005)] as drivers change speed in response to congestion, weather conditions and the behavior of other drivers, as well as the number of available lanes change due to weather conditions, traffic issues and other events. Accurately estimating road capacity from sparse and noisy observations of traffic speed and density at points in the traffic network is a significant challenge and improving on these estimates will lead to better travel time forecasts.

Our approach develops a particle filtering and learning algorithm for estimating road capacity. We build on existing estimation methods in a number of ways:

1. Incorporation of sequential parameter learning in order to update the model in real time.
2. A predictive likelihood particle filter that provides an efficient estimation strategy and is less sensitive to measurement outliers.

We apply our methodology to traffic flow data from Chicago's interstate I-55 highway and show how parameter learning effectively handles a dynamic environment, including shock waves. Bayesian learning, which is central to our methodology, corrects for bias that results from estimation with fixed parameters. We also show that our algorithm identifies the drop in road throughput as a result of an accident.

Particle filtering allows for posterior estimation of the most recent state. The low computational complexity of particle filtering makes frequent updating feasible, whereas MCMC's computational cost grows linearly with the length of the data. For previous MCMC applications in transportation, see Tebaldi and West (1998) for inferring network route flows, and Westgate et al. (2013) for travel time reliability for ambulances using noisy GPS for both path travel time and individual road segment travel time distributions. Anacleto, Queen and Albers (2013) develop a dynamic Bayesian network to model external intervention techniques to accommodate situations with suddenly changing traffic variables. Chiou, Lan and Tseng (2014) provide a nonparametric prediction model for traffic flow trajectories, and Chiou (2012) proposes using a functional mixture prediction approach.

Previous work on estimating traffic flows use extensions of the Kalman filter and rely heavily on Gaussianity assumptions; see Gazis and Knapp (1971), Schreiter et al. (2010), Wang and Papageorgiou (2005), Work et al. (2008). Sun, Muñoz and Horowitz (2003) considered mixture Kalman filters for traffic state estimation in the context of ramp metering control. Particle filters have previously been applied to traffic flow problems; see Mihaylova, Boel and Hegyi (2007) who use the evolution dynamics as a proposal distribution before resampling, the so-called bootstrap or sampling/importance resampling (SIR) filter. We improve the efficiency

for inference and prediction with a fully adopted filter and our approach naturally incorporates particle learning. We build on existing work on parameter learning in transportation. For example, Dervisoglu et al. (2009) develop a quantile regression methodology that re-estimates parameters every five minutes based on traffic flow and density measurements. Wang and Papageorgiou (2005) propose an extended Kalman filter with boundary condition estimation. The advantage of particle filtering over traditional Kalman filtering is the ability to handle nonnormal posterior distributions that result, for example, from nonlinearity. Section 2.6 shows that the distribution of uncertainty about state is a mixture at some points in time and this leads us to use particle filters, that do not rely on normality assumption.

Real-time estimation and short-run prediction of traffic conditions play a key role in Intelligent Transportation Systems (ITS). Current Vehicle Navigation Systems and Traffic Management Systems use forecasts of traffic flow variables, such as traffic volume, travel speed or traffic density ranging from 5–30 minutes ahead. There are a number of real-world applications:

Advanced Traveler Information Services (ATIS). Multiple studies have shown the positive impacts of providing information on traffic flow conditions to the public [Chorus, Molin and Van Wee (2006)], as it can potentially lead to congestion relief [Arnott, De Palma and Lindsey (1991)]. Travel information is provided in multiple ways, for example, by transportation system managers such as local departments of transportation via variable message signs or radio, automakers through in-dash navigation, technology companies through phone apps or web, fleet managers and transit operators.

Transportation Planning. Benefits of Intelligent Transportation Systems are studied by local governments based on system performance data before and after ITS is deployed. An accurate comparison of the benefits to travel times requires efficient estimation of the network states.

Control of Transportation Operation. For traffic control applications, we need to efficiently estimate the formation of traffic congestion. Accurate knowledge of the current state allows transportation system managers to provide a reasonable forecast of traffic conditions and to improve traffic flows using such techniques as ramp metering and speed harmonization.

The rest of the paper proceeds as follows. Section 2 develops a statistical treatment for the LWR model by representing it as a nonlinear state-space model. The key input to the LWR model, the fundamental diagram (or flux function), which links traffic flow and density is discussed. The parameters of the fundamental diagram need to be estimated in an online fashion. Section 3 provides a particle filtering algorithm for inference and prediction that provides online real-time inference for fundamental diagram parameters and traffic density state. Section 4 illustrates how our methodology can learn road capacity when applied to data measured during a major highway accident. Section 5 illustrates our methodology with a simulation study of rush hour traffic on Chicago's I-55. Finally, Section 6 concludes with directions for future research.



FIG. 1. Locations of the Loop Detectors in Chicago.

2. LWR traffic flow model.

2.1. Model and data description. Traffic flow data is available from the Illinois Department of Transportation; see Lake Michigan Interstate Gateway Alliance (<http://www.travelmidwest.com/>), formally the Gary–Chicago–Milwaukee Corridor (GCM). The data is measured by loop-detector sensors installed on interstate highways. Loop-detector is a simple presence sensor that measures when a vehicle is present and generates an on/off signal. There are over 900 loop-detector sensors that cover the Chicago metropolitan area. Figure 1 illustrates the locations of the detectors in the region. Since 2008, Argonne National Laboratory has been archiving traffic flow data every five minutes from the grid of sensors. Data contains averaged *speed*, *flow* and *occupancy*. Occupancy is defined as percent of time a point on the road is occupied by a vehicle and flow is the number of off-on switches. Illinois uses a single loop-detector setting and speed is estimated based on the assumption of an average vehicle length.

2.2. Traffic flow parameters. The primary variable of interest is traffic density, which is a macroscopic characteristic of traffic flow and the control variable of interest in transportation system management strategies. Traffic density is defined as a number of vehicles per unit of length. Densities vary between zero and jam density which corresponds to vehicles being bumper-to-bumper. Typically, jam density value is around 1 vehicle per 6.5 meters per lane. Another important value related to density is the critical density, denoted by the density level at which the maximum flow (throughput) is achieved. The maximum flow measured in vehicles

per unit of time is called capacity, it is typically achieved at density level of around 1 vehicle per 32 meters.

It is natural to divide the flow regimes roughly into two subcategories. Density values up to 1 vehicle per 32 meters correspond to a free-flow regime, when there are no interactions between the vehicles, and vehicles travel at the desired speed. The second regime corresponds to densities above 1 vehicle per 32 meters; at roughly this density vehicles start interacting with each other and that leads to slow downs and flow reduction.

Our observed data that comes from a presence sensor is occupancy rather than density. Occupancy is defined as percentage of time a point on a road segment was occupied by a vehicle, thus it varies between 0 (empty road) and 100 (complete stand still). Assuming the average vehicle length does not vary over time, the density and occupancy are related through a simple linear transformation [May (1990)]. Throughout the paper we assume a constant vehicle length for every sensor in the region and treat density and occupancy interchangeably.

Two other macroscopic traffic flow parameters, namely. Speed and flow, are related through the following relation:

$$(2.1) \quad v(x, t) = \frac{q(x, t)}{\rho(x, t)},$$

where v = speed (miles per hour), q = flow (vehicles per hour), ρ = density (vehicles per lane-mile).

The three traffic flow parameters can change over space and time. Tracking these flow parameters can be particularly challenging due to discontinuities in them that are called shock waves. A shock wave can be a platoon of vehicles moving on an otherwise empty road, thus we have a nonzero density propagating in time and space. In other cases, the shock wave corresponds to a change in the flow regime, when fast-moving vehicles reach the end of a congestion queue and need to abruptly slow down, or vice versa, when we have queue dissipation and vehicles leave a bottleneck and can revert to the desired travel speed.

2.3. The LWR model and fundamental diagram. In Section 2.2 we considered traffic flow as a function of location x and time t . The flow-density relation, which is called the *fundamental diagram*, allows us to calculate flow via density $q(x, t) = q(\rho(x, t))$. The LWR model is a macroscopic traffic flow model. It is a combination of a conservation law defined via a partial differential equation and a fundamental diagram. The nonlinear first-order partial differential equation describes the aggregate behavior of drivers. The density $\rho(x, t)$ and flow $q(x, t)$, which are continuous scalar functions, satisfy the equation

$$(2.2) \quad \frac{\partial \rho(x, t)}{\partial t} + \frac{\partial q(x, t)}{\partial x} = 0.$$

Derivation of the model is presented in Appendix A. This equation can be solved numerically by discretizing time and space. In its simplest form, imagine a homogeneous road segment (with no change in number of lanes and no intersections) cut into M cells. Let ρ_i be the density in cell i (in veh/m) and q_i the exit flow of cell i (in veh/s). For a road segment, with given boundary conditions, the LWR computes the conditions inside the domain. Boundary conditions can be either measured by fixed sensors such as loop detectors or estimated from GPS probe data based as shown by Claudel and Bayen (2010). Statistical inference is required to update the missing states, learn the parameters of interest and predict forward using the dynamics of the LWR model, based on noise and possibly partially measured boundary conditions.

An important feature of the LWR model is the emergence of a shock wave of traffic due to the density-dependent local propagation velocities. The fundamental diagram is central to its specification. The diagram describes a functional relation between flow and density. For example, Figure 2(a) illustrates empirical data of volume versus occupancy. The theoretical form of the so-called triangular fundamental diagram is shown in Figure 2(b). It has two velocities of density variations: one for free-flow traffic (green) and one for congested traffic (red). This specification allows for an efficient Godunov scheme, to solve the nonlinear evolution dynamics. We need to provide the model with an accurate assessment of the current density state vector and the parameters of the fundamental diagram. The *fundamental diagram* is a key input into the specification of the LWR model, which expresses the relationship between traffic density and flow. Figure 2 motivates the choice of a so-called triangular diagram by showing the empirical flow and occupancy for a highway segment in the Chicago metropolitan area.

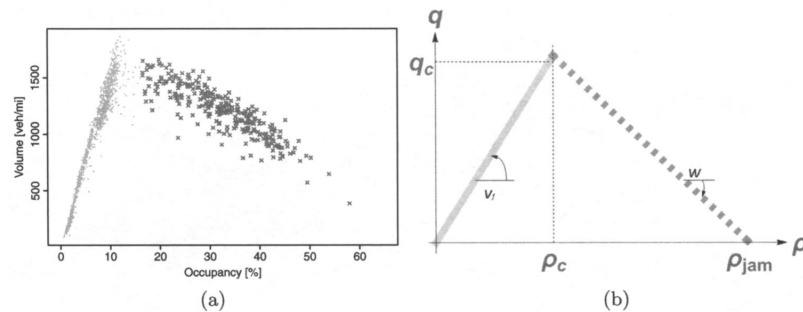


FIG. 2. Fundamental diagram. (a) Measured occupancy-flow relation. (b) Triangular fundamental diagram. The left panel (a) shows the occupancy-flow relation based on measured data on I-55 North Bound. The right panel (b) shows theoretical shape of the fundamental and the parameters that describe the diagram. On both panels the left part of the diagram [triangles in (a) and solid line in (b)] describes the density variations for free-flow traffic and the right part [crosses in (a) and dashed line in (b)] describes congested traffic.

Throughout our analysis we assume a homogeneous road segment with a fundamental diagram that does not depend on time and space. By homogeneous, we mean that the road segment has homogeneous width and number of lanes and no intersections or traffic merge/diverge sections.

The analytical formula for the triangular fundamental diagram is as follows:

$$(2.3) \quad q(\rho) = \begin{cases} \frac{q_c}{\rho_c} \rho, & \rho < \rho_c, \\ q_c \frac{\rho_{\text{jam}} - \rho}{\rho_{\text{jam}} - \rho_c}, & \rho \geq \rho_c, \end{cases}$$

where q_c = representing the critical flow (capacity), ρ_c = critical density, ρ_{jam} = jam density. We denote the set of three parameters by $\phi = (q_c, \rho_c, \rho_{\text{jam}})$.

The velocity of a shock wave propagation on a road segment can be calculated using the fundamental diagram parameters, via the Rankine–Hugoniot relation [LeVeque (2002)]. It determines the shock wave velocity as the velocity of the shock w times the jump in density which equals the jump in flow in the two regions separated by the shock where

$$(2.4) \quad w = \frac{q(\rho_l) - q(\rho_r)}{\rho_l - \rho_r} \quad \text{and} \quad v_f = \frac{q_c}{\rho_c}.$$

The direction of the shock wave propagation depends on the sign of $q(\rho_l) - q(\rho_r)$. Here v_f is a free-flow speed on a link and $q_c = \max_{\rho} q(\rho)$ is the critical flow or capacity of the link. Correspondingly, $\rho_c = \arg \max_{\rho} q(\rho)$ is called the critical density. The pair (q_c, ρ_c) is the traffic flow breakdown point for a road segment.

Calibrating the model parameters can be done in a number of ways. The standard approach uses values from the Highway Capacity Manual [Transportation Research Board (2010)] that provides a look-up table for road capacity based on road type and number of lanes.

However, in practice, the parameters are not fixed and change over time. To empirically illustrate the stochastic nature of the parameters, we estimate capacity and critical density from the measurements for 242 days in 2009, on a segment of interstate highway I-55 in Chicago. Holidays and weekends as well as days with unreliable measurements were excluded. Figure 3 plots ρ_c and q_c across the days. Clearly, there is a linear relation between q_c and ρ_c . Road capacity can vary from day to day and its distribution has a heavy left tail. On the other hand (for our data set), the critical density ρ_c has a relatively tight distribution around the value 0.023 veh/m.

This nonstatic nature of the parameters motivates the need for a sequential on-line parameter learning algorithm.

2.4. Traffic flow dynamics as a nonlinear state-space model. Let y_t denote the observed traffic density data and $y^t = (y_1, \dots, y_t)$ be the current history of data. Let θ_t be a hidden state vector of traffic densities. We assume that boundary

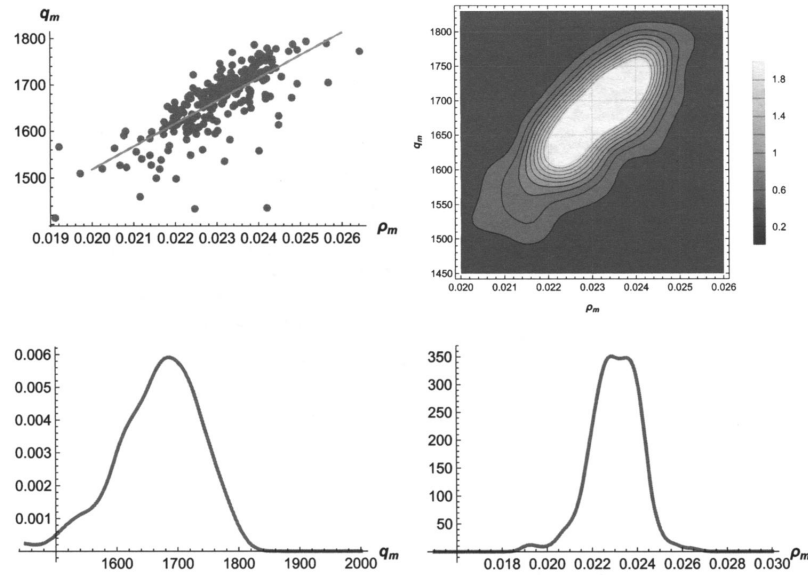


FIG. 3. Joint and marginal distributions for critical flow and density.

conditions ρ_{0t} and $\rho_{(M+1)t}$, which represent traffic states on the downstream and upstream ends of a road segment, are given, as well as an initial condition θ_0 . In practice, boundary conditions are measure from sensors, such as loop detectors and radars, available at both ends of a road segment, and initial conditions either assume an empty road or state of traffic measured from cameras or satellites.

We denote

$$\theta_t = (\rho_{1t}, \dots, \rho_{Mt}).$$

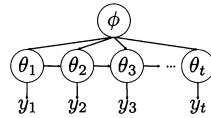
The expectation conditional of the next state $E(\theta_{t+1}|\theta_t) = f_\phi(\theta_t)$ is given by the solution of the LWR model. Here ϕ denotes unknown parameters. A numerical Godunov scheme computes $f_\phi(\theta_t)$ given the parameters, ϕ , of the triangular fundamental diagram.

Our model has a state-space formulation of an observation and evolution system given by

$$(2.5) \quad \text{Observation:} \quad y_{t+1} = H_{t+1}\theta_{t+1} + \varepsilon_{t+1}^v; \varepsilon_{t+1}^v \sim N(0, V_{t+1}),$$

$$(2.6) \quad \text{Evolution:} \quad \theta_{t+1} = f_\phi(\theta_t) + \varepsilon_{t+1}^w; \varepsilon_{t+1}^w \sim N(0, W_{t+1}),$$

where V_t and W_t are evolution and equation error, respectively, and y_{t+1} = vector of measured traffic flow density, f_ϕ = LWR evolution equation calculated via Go-

FIG. 4. *Parameter learning graphical model.*

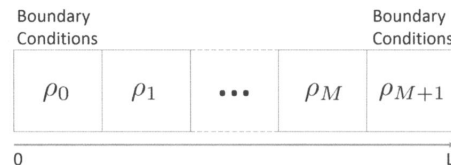
dunov's schema, $\phi = (q_c, \rho_c, \rho_{\text{jam}})$ triangular fundamental diagram parameters. The observation matrix H_{t+1} picks out cells with measurements available.

Figure 4 provides the graphical model for the state evolution structure. Our goal is to develop a particle filter to draw samples from the filtered posteriors $p(\theta_t|y^t)$ and $p(\phi|y^t)$. The operator $H_t: \mathbb{R}^M \rightarrow \mathbb{R}^k$ is the measurement model that depends on the sensor type, and in our setting we make it linear. In a simplest case $H_t = H$ is a projection operator, which “removes” nonmeasured elements from the state vector.

While we treat the parameters ϕ of the LWR model as static, our model can easily be extended to allow for stochastic evolutions or characteristics to govern the dynamics.

2.5. Godunov's scheme. The LWR model (2.2) describes the evolution of traffic flow on a road segment with uniform topology, as shown in Figure 5 (see Appendix A). The change in road segment characteristics (crossing, number of lanes, speed limit, curvature, etc.) can be modeled using a junction. The treatment of junctions requires specific efforts for physical consistency and mathematical compatibility with the link model. For uniqueness of the solution of the junction problem, different conditions have been used, for instance, maximizing the incoming flow through the junction was suggested by Daganzo (1995) and Coclite, Garavello and Piccoli (2005). Holden and Risebro (1995) consider maximizing a concave function of the incoming flow. A formulation using internal dynamics for the junction is equivalent [Lebacque (2005)] to the vertex models for the merge and diverge junction; see Garavello and Piccoli (2006) for more details.

Standard finite difference schemes are too inaccurate for solving the LWR model; Godunov (1959) showed that a first order finite difference scheme is inaccurate for calculating with a small time step. Moreover, none of the second order difference schemes preserve monotonicity of the ρ_0 , and thus are not applicable.

FIG. 5. *Underlying state space for a road segment.*

Given an initial condition $\rho_0(x)$, $x \in [0, L]$, propagating the LWR model requires solving the associated Cauchy problem. If the initial condition is piecewise constant (which is the case for many numerical approximations) and self-similar, this reduces to a Riemann problem. Godunov's scheme then solves a Riemann problem between each cell. This is an initial value problem with initial conditions having a single discontinuity

$$(2.7) \quad \rho_0(x) = \begin{cases} \rho_l, & x < 0, \\ \rho_r, & x > 0. \end{cases}$$

For the Riemann problem, the speed of the shock wave propagation is given by the Rankine–Hugoniot relation (2.4). Heuristically, imagine at initial time $t = 0$ that there are two regions in the domain with different values of thermodynamic parameters (flow, density and speed in our case). The two regions are divided by a thin membrane and at the initial time the membrane is removed. The computational problem is to find the values of thermodynamic parameters at all future times.

According to Godunov's scheme, we calculate the iterates

$$(2.8) \quad \rho_i^{n+1} = \rho_i^n + \frac{\tau}{h} (q_G(\rho_{i-1}^n, \rho_i^n) - q_G(\rho_i^n, \rho_{i+1}^n)),$$

where ρ_i^n is the density value at the point with coordinates $x = ih$, $t = n\tau$, with h a space discretization step and τ a time discretization step.

The function $q_G(\rho_l, \rho_r)$ is defined by

$$(2.9) \quad q_G(\rho_l, \rho_r) = \begin{cases} q(\rho_l), & \rho_r < \rho_l \leq \rho_c, \\ q(\rho_c), & \rho_r \leq \rho_c \leq \rho_l, \\ q(\rho_r), & \rho_c \leq \rho_r < \rho_l, \\ \min(q(\rho_l), q(\rho_r)), & \rho_l < \rho_r. \end{cases}$$

Typically, a virtual cell is introduced on both sides of the domain to include boundary conditions (in and out flow). This leads to a left boundary

$$\rho_0^{n+1} = \rho_0^n + \frac{\tau}{h} (q_G(\rho_{-1}^n, \rho_0^n) - q_G(\rho_0^n, \rho_1^n)),$$

$$\text{with } \rho_{-1}^n = \frac{1}{\tau} \int_{(n-1/2)\tau}^{(n+1/2)\tau} \rho(0, t) dt,$$

and right boundary

$$\rho_M^{n+1} = \rho_M^n + \frac{\tau}{h} (q_G(\rho_{M-1}^n, \rho_M^n) - q_G(\rho_M^n, \rho_{M+1}^n)),$$

$$\text{with } \rho_{M+1}^n = \frac{1}{\tau} \int_{(n-1/2)\tau}^{(n+1/2)\tau} \rho(L, t) dt.$$

Numerical stability in space and time is ensured by the Courant–Friedrichs–Lewy type condition [Courant, Friedrichs and Lewy (1928)]: $\tau \leq h/|v_{\max}|$, where v_{\max} is the maximum wave velocity present in the meshed domain at any given point in time.

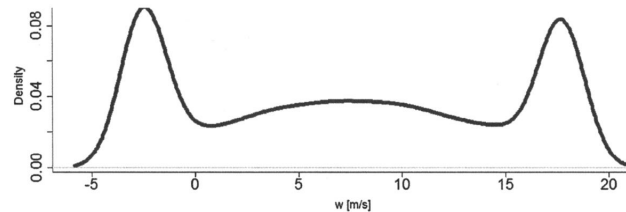


FIG. 6. Uncertainty of shock wave propagation speed.

2.6. State uncertainty is a mixture distribution. When uncertainty about the traffic state gets propagated from one time step to another using Godunov's scheme, the current unimodal distribution can update to a mixture distribution. For example, this happens at the location of a shock wave, when the cell on the right is in a free-flow regime and the cell on the left is in a congested regime. This can be demonstrated by a simple Monte Carlo experiment. Consider two consecutive cells, with densities ρ_l and ρ_r correspondingly, both following a truncated normal distribution. Assume $\rho_l \sim TN(\mu = 0.02, \sigma = 0.01, a = 0, b = 0.2)$ and $\rho_r \sim TN(\mu = 0.03, \sigma = 0.01, a = 0, b = 0.2)$, where a and b are lower and upper bounds of a truncated normal distribution correspondingly. Using a triangular fundamental diagram with $q_c = 1600$ veh/h, $\rho_c = 0.025$ veh/m, and $\rho_{jam} = 0.2$ veh/m, we can calculate the speed of the shock wave propagation w given by equation 2.4. We then simulate the distribution over w , using $N = 1000$ samples. Figure 6 shows the results of the experiment. The uncertainty over speed propagation is a bimodal mixture distribution, implying the uncertainty about the density at the future times is also a mixture. Our example in Section 5 shows that the behavior of uncertainty about traffic flow density state matches this bimodal shape found here.

3. Particle filtering of the LWR model.

3.1. A fully adapted particle filter. Particle filtering methods are designed to provide sequential state inference from the set of filtered posteriors $p(\theta_t | y^t)$; see, for example, Carpenter, Clifford and Fearnhead (1999), Carvalho et al. (2010), Gordon, Salmond and Smith (1993), Liu and West (2001), Pitt and Shephard (1999), Storvik (2002). Our algorithm will be based on the Liu and West (2001) filter. The major difference is a fully adapted filter that resamples first using the predictive distribution and propagates forward using the conditional posterior. The fully adapted filter mitigates particle filter degeneracy, although the usual compounding of the Monte Carlo errors still exists [Godsill, Doucet and West (2004)].

The predictive likelihood for the next observation, y_{t+1} , is required to implement our particle filter. Given the current state variable θ_t , the predictive likelihood

is defined by

$$p(y_{t+1}|\theta_t, \phi) = \int p(y_{t+1}|\theta_{t+1}, \phi) p(\theta_{t+1}|\theta_t, \phi) d\theta_{t+1}.$$

Propagation of states requires the conditional posterior for the next state $p(\theta_{t+1}|\theta_t, \phi, y_{t+1})$. This density can be computed using the model assumptions via the system of distributions

$$\begin{aligned} p(y_{t+1}|\theta_{t+1}, \phi) &\sim N(H_{t+1}\theta_{t+1}, V_{t+1}), \\ p(\theta_{t+1}|\theta_t, \phi) &\sim N(f_\phi(\theta_t), W_{t+1}). \end{aligned}$$

Marginalizing out θ_{t+1} leads us to distributions

$$p(y_{t+1}|\theta_t, \phi) \sim N(H_{t+1}f_\phi(\theta_t), H_{t+1}W_{t+1}H_{t+1}^T + V_{t+1}).$$

For propagation of θ_{t+1} , we use Bayes' rule and the conditional posterior

$$p(\theta_{t+1}|\theta_t, \phi, y_{t+1}) \sim N(\mu_{t+1}, C_{t+1}),$$

where the mean and variance (μ_{t+1}, C_{t+1}) follow the Kalman recursion [Doucet, Godsill and Andrieu (2000)]:

$$\begin{aligned} \text{Forecast:} \quad \mu_f &= f_\phi(\theta_t), \quad C_f = W_{t+1}, \\ \text{Kalman Gain:} \quad K &= C_f H_{t+1}^T (H_{t+1} C_f H_{t+1}^T + V_{t+1})^{-1}, \\ \text{Measurement Assimilation:} \quad \mu_{t+1} &= \mu_f + K(y_{t+1} - H_{t+1}\mu_f), \\ C_{t+1} &= (I - K H_{t+1}) C_f. \end{aligned}$$

To develop our particle filter, we now factorize the joint conditional distribution as

$$p(y_{t+1}, \theta_{t+1}|\theta_t, \phi) = p(y_{t+1}|\theta_t, \phi) p(\theta_{t+1}|\theta_t, \phi, y_{t+1}).$$

The goal is to obtain the new filtering distribution $p(\theta_{t+1}|y^{t+1})$ from the current $p(\theta_t|y^t)$ and to provide a particle approximation to the parameter posterior, $p(\phi|y^t)$. We start with a particle (a.k.a. random histogram of draws) filtering approximation to the joint distribution of the state and parameters, denoted by

$$p^N(\theta_t, \phi|y^t) = \frac{1}{N} \sum_{i=1}^N \delta_{(\theta_t, \phi)}^{(i)},$$

where δ is a Dirac measure. As the number of particles increases $N \rightarrow \infty$, the law of large numbers guarantees that this distribution converges to the true filtered distribution $p(\theta_t, \phi|y^t)$.

For the next marginal posterior distribution, the Bayes rule yields

$$p^N(\theta_{t+1}|y^{t+1}) = \sum_{i=1}^N w_t^{(i)} p(\theta_{t+1}|\theta_t, \phi)^{(i)}, y_{t+1}),$$

where the particle weights are determined by

$$w_t^{(i)} = \frac{p(y_{t+1} | (\theta_t, \phi)^{(i)})}{\sum_{i=1}^N p(y_{t+1} | (\theta_t, \phi)^{(i)})}.$$

The algorithm consists of three steps:

Step 1. (Resample) Draw an index $k(i) \sim \text{Mult}_N(w_t^{(1)}, \dots, w_t^{(N)})$ for $i = 1, \dots, N$.

Step 2. (Propagate) Draw $\theta_{t+1}^{(i)} \sim p(\theta_{t+1} | (\theta_t, \phi)^{k(i)}, y_{t+1})$ for $i = 1, \dots, N$.

Step 3. (Replenish) Draw $\phi^{(i)} \sim \frac{1}{N} \sum_{i=1}^N \delta_{[-\varepsilon, \varepsilon]}(\phi^{k(i)})$,

where $\delta_{[-\varepsilon, \varepsilon]}(\cdot)$ denotes the Dirac measure in an interval $[-\varepsilon, \varepsilon]$. Thus, we resample $\phi^{k(i)}$ from mixture uniform distribution with support $[\phi^{k(i)} - \varepsilon, \phi^{k(i)} + \varepsilon]$, $i = 1, \dots, N$, and equal mixing rights. The jittering parameter ε is used to calculate unique ϕ^i particles. Both $\theta_{t+1}^{(i)}$ in step 2 and $\phi^{(i)}$ in step 3 are drawn based on resampled $\phi^{k(i)}$, thus the resampling creates a new set of particles $(\theta_t, \phi)^{k(i)}$. Steps 1 and 2 of the algorithm were suggested in the auxiliary particle filter of Pitt and Shephard (1999).

It has been previously shown that particle filters suffer the degeneracy issue when the number of particles is not sufficient [Bengtsson, Bickel and Li (2008), Snyder (2011)]. However, our approach relies on predictive likelihood and is less prone to a degeneracy issue, which plagues standard sample-importance resample filters.

4. Real-time accident modeling. We illustrate our methodology on a data set from an accident on I-55. We show how quickly our approach can identify a drop in capacity (critical flow) due to an accident. On May 9, 2014, a semi-tractor trailer caught fire (CBS Chicago) at 6:40 AM on interstate highway I-55 near Weber Road in Romeoville, Illinois, which is a southwest suburb of Chicago. The police shut down the southbound lanes. As is commonplace, the accident was visible from the other side of the road, and the “rubbernecking” effect, drivers slowing down to watch an accident, caused a dramatic reduction in capacity and congestion.

There are several reasons for capacity reduction during an accident. Under normal conditions, an average time delay before a vehicle starts accelerating following a leader is half a second. However, during an accident there is a large difference in times that drivers took to look at the accident location before accelerating. These results were obtained by Knoop, Hoogendoorn and Van Zuylen (2008) via analyzed video taken by helicopter from accident locations. Most of the vehicles would accelerate at the usual rate out of the jam, and the shock wave would move backward. However, it only requires a small fraction of drivers that keep driving slowly until they reach the accident location to cause large escape times at the location of the incident and hence for the low capacity. There is also heterogeneity



FIG. 7. Accident location. (a) Accident and loop detector locations. (b) Image of the accident from the roadside camera. The left panel (a) shows the satellite image with the location of the accident identified (red rectangle) and two loop detectors located (green circles) on the opposite direction before and after the accident location (credit: Bing Maps). The right panel (b) shows the image of the truck on fire taken by Illinois Department of Transportation's roadside camera on the day of the accident.

in acceleration delays between left lane (closest to accident) and right lane. On average, in the left lane, cars take longer before accelerating.

Figure 7(a) shows the map location of the accident. Figure 7(b) shows two of the loop detectors located before and after the accident location from which the data was collected.

The length of the road segment between two loop detectors is 845 meters and we discretized it with four cells, with each space step $h = 845/4 = 211$ meters and used time step $\tau = 5$ minutes. This combination of time and space step satisfies the Courant–Friedrichs–Lewy condition so that numerical stability in space and time is ensured. Further, a five minute interval was chosen since it is a standard interval over which the measured data is averaged to provide smooth input data. Our initial prior on road capacity is assumed to be uniform, with $q_m \sim U[1440, 1560]$ veh/h, and we set critical density to $\rho_m = 0.025$ veh/m; both are based on empirical observations of typical ranges for those parameters as shown in Figure 3. To replenish the parameters (step 3 of the algorithm), we used $\varepsilon_{q_m} = 50$ veh/h for capacity and $\varepsilon_{\rho_m} = 0$ veh/m for critical density, as there is no learning for this parameter. The value for ε_{q_m} is based on empirical observations, that capacity change usually does not exceed 50 veh/h within a five minute interval.

We have chosen the measurement noise's standard deviation to be 0.2×10^{-2} veh/m, and standard deviation for the evolution equation error to be equal to 0.1×10^{-2} veh/m. Given that we did not have access to manufacturer's specifications of the loop detectors, we use a value within the guidelines of the specification. The error for the evolution equation was chosen to be consistent with the results reported in Chu et al. (2011), where authors report that standard deviation of the LWR model forecast error is usually under 3%, but higher for congested flow

when compared with observation data from motorway sensors. In our numerical example, we use 4%.

To address the problem of model identification, we utilize the relation between free-flow speed, capacity and critical density, namely, $v_f = q_c / \rho_c$. Based on the data measured on a typical day during off-peak hours, the free-flow speed $v_f \approx 17$ m/s. Our particle weights are regularized by

$$w_t^{(i)} = \frac{p(y_{t+1} | (\theta_t, \phi)^{(i)}) \varphi(q_c^{(i)} / \rho_c^{(i)}, v_f, \sigma_{v_f})}{\sum_{i=1}^N [p(y_{t+1} | (\theta_t, \phi)^{(i)}) \varphi(q_c^{(i)} / \rho_c^{(i)}, v_f, \sigma_{v_f})]},$$

where φ is the p.d.f. of the normally distributed variable. The prior error standard deviation was set at $\sigma_{v_f} = 5$ m/s. Choice of both v_f and σ_{v_f} is based on empirical observations.

Figure 8(a) compares the road capacity learned by the algorithm on the day of the accident and from the previous day, which was accident free, with similar weather conditions. Figure 8(b) shows the measured speed by the south loop detector on both days. There is a time lag of approximately 15 minutes between traffic flow speed reverts to a normal level and capacity recovers. This time lag corresponds to three measurements (data is reported every five minutes) and is explained by the time it takes the algorithm to learn the capacity.

Our algorithm captures the effect of capacity degradation as a result of the accident. We provide 95% Bayes credible intervals to demonstrate that uncertainty about the estimate is larger during the normal operating mode and lower during the periods of capacity degradation and recovery. If we compare the speed plot and capacity plot in Figure 8(a) and (b), we can see that the slope of the speed

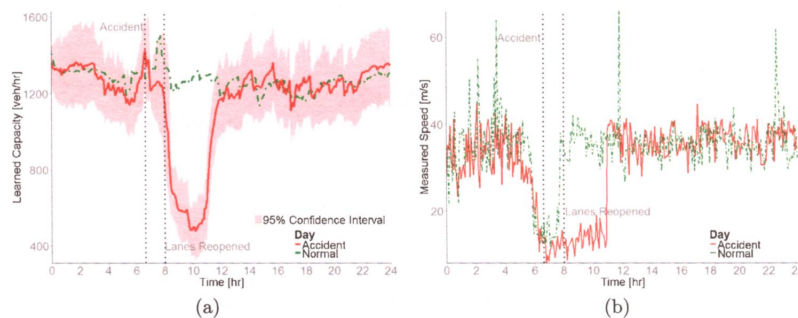


FIG. 8. Comparison of the learned capacity and measured speed on Thursday, May 8th (normal day) and Friday, May 9th (accident day). (a) Learned capacity of the road segment. (b) Measured speed at the south loop detector. On both plots the left vertical line identifies the time when accident happened (6:40 AM) and the second vertical line corresponds to the time when all of the lanes reopened at 8:00 AM, according to the news report. The number of particles was chosen $N = 5000$. Accident data was plotted using a solid line and normal day with a dashed line.

curve on an accident day is much steeper than the slope of the learned capacity curve; it is due to the fact that there is some delay associated with the learning process. In other words, the algorithm does not learn that the flow regime has recovered instantaneously, but rather it takes three to five measurements before it learns.

Under normal conditions, a full-width freeway lane has a capacity of 2000 passenger vehicles per hour [Transportation Research Board (2010)], with a truck being counted as 1.5 passenger cars. In Illinois, the loop detectors give reliable data for the vehicle counts but not for different vehicle classes and it is hard to identify the share of trucks in the traffic flow; consequently, the flow is measured in vehicles per hour and not in passenger car units per hour. Thus, the learned capacity is around 1500 veh/h on a normal day, that is consistent with the theoretical estimate from the highway capacity manual. On the accident day we detect a reduction in capacity of up to 66%. This is similar to the results of Knoop, Hoogendoorn and Van Zuylen (2008) who use helicopter images from Netherlands roads to observe a 50% reduction of capacity, due to the reduction of the discharge rate at the bottleneck (accident location) due to rubbernecking. A larger drop in our case might be explained by regional differences in driving style. American drivers might be driving more carefully in the presence of an accident, and by the fact that a truck on fire is more “spectacular” than a regular vehicle crash, with people spending more time to observe. Such a drop in the flow rates is remarkable given the absence of any physical obstacles.

5. Calibration experiment. The previous example illustrates a drop in capacity due to an accident. However, we do not know if the drop size is properly estimated since the ground truth is unobserved. To demonstrate that our algorithm properly captures state and parameter dynamics, we must use simulated data with a realistic traffic flow pattern. We simulate data that mimics traffic flow on Chicago’s I-55 highway. Figure 9 below shows traffic patterns on February 6th, 2009 (Friday) and all five work days of the following week (week of February 9th). Several conclusions can be drawn from the traffic patterns:

- (i) Break down start times are different from day to day, even on the same day of the week (Fridays) of different weeks.
- (ii) The duration of the flow at the lowest speed is different, with Wednesday being the worst and Thursday the best.
- (iii) The breakdown period is shorter than the recovery period.

The latter point follows from the asymmetric shape of the triangular fundamental diagram, where the free-flow speed v_f (speed at which drivers arrive to the end of the congestion queue) is higher than the backward wave propagation speed w (speed at which drivers depart from the front of the congestion queue).

Our road segment is 1.5 kilometers long and we choose a time horizon of 1600 seconds. Figure 10 shows our road segment model and its discretization scheme

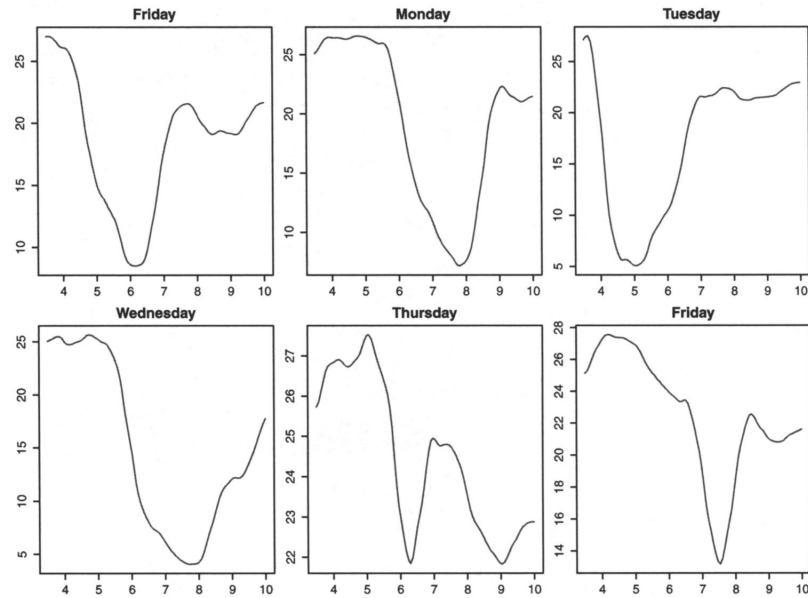


FIG. 9. Chicago I-55 Work Day Morning Peak Traffic Patterns. The x-axis corresponds to the time of the day.

with five internal cells and 2 boundary cells. Our discretization grid cell is of length $h = 300$ meters and time interval of $\tau = 5$ seconds. The initial conditions are set to be uniform traffic density of 0.01 veh/m.

To mimic a typical morning commute pattern, we have chosen boundary conditions so that our simulated data set begins with a free-flow traffic regime followed

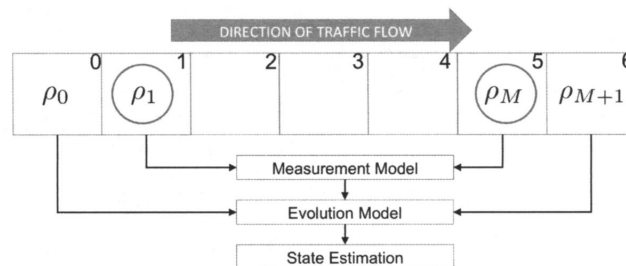


FIG. 10. Simulated stretch of a freeway. The arrow shows the direction of the traffic flow. It is assumed that measurements from circled first and last cells of the domain are available.

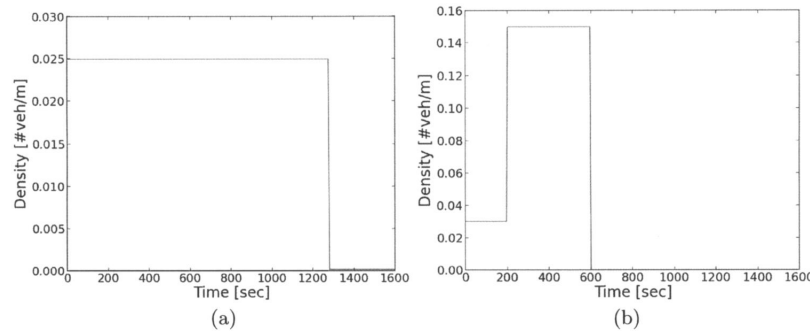


FIG. 11. Boundary conditions used to produce simulated data. (a) Left (cell 0). (b) Right (cell 6).

by a breakdown and then recovery. The breakdown starts 3 minutes into the simulation and the recovery starts at the 10 minute mark. In the left boundary cell (cell 0) we have a constant vehicle density followed by a drop in density to zero; see Figure 11(a). It mimics the constant inflow of morning commuters that eventually stops. On the right boundary cell (cell 6) we have uncontested density at the beginning, followed by density of 0.145 veh/m, which represents heavily congested traffic flow, and followed by a drop in density to zero; see Figure 11(b). The right boundary condition corresponds to a location where an on-ramp merges into a highway. When the flow on the ramp is high, a bottleneck is created at the merge location. The boundary conditions are shown in Figure 11. Over the course of simulation we changed the traffic flow parameters. Capacity and critical density parameters used to produce simulated data are shown in Figure 13.

To simulate the measured data, we compute the solution of the LWR model for the measurement cells 1 and 5 and add noise to it. We have chosen the noise standard deviation to be 0.8×10^{-2} veh/m, and standard deviation for the evolution equation error to be equal to 0.1×10^{-2} veh/m.

Figure 12 compares the estimated traffic density in cell 3 with the true simulated traffic density for two different scenarios. In the first scenario we used the parameter learning step of the algorithm and in the second scenario we kept capacity and critical density parameters fixed. We can see the sensitivity to parameter learning. Without learning, the density profile is shifted in a meaningful way. Clearly, full Bayesian parameter learning corrects this bias.

To illustrate the dynamics of parameter learning, we change the LWR parameters ρ_c and q_c several times throughout the simulation, as mentioned above. In principle, we could also directly model ϕ_t with its own state evolution.

Figure 13 shows the expected value and 95th percentile of the filtered posterior distribution of the model parameters. We can see, as expected, there is a certain delay between the underlying parameter change and the filtering algorithm cap-

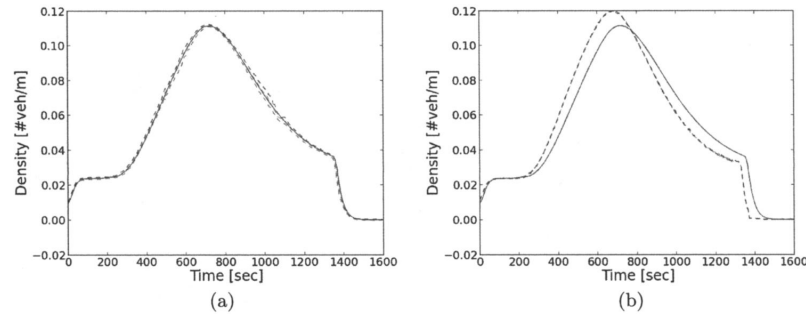


FIG. 12. *Estimated density at cell 3. (a) With parameter learning. (b) Without parameter learning. The blue line on both plots is ground truth and the dotted line is the filtered value computed by the algorithm. The number of particles was chosen $N = 1000$.*

tures the change. Change in capacity is picked up faster than change in critical density.

In Section 2.6 we showed that the distribution over traffic density is a mixture distribution at the locations when the density in the left cell is below critical density and the density in the right cell is above. To further demonstrate this fact, Figure 14 shows the distribution over density at cell 3 before and after the shock wave travels through the cell.

Figure 12 shows that a shock wave travels through cell 3 between $t = 600$ and $t = 700$. We can see that the distribution over state is unimodal at those time steps. However, in between, it is a mixture.

Appendix B develops the projection operator H and the necessary Kalman recursions for this and the following examples.

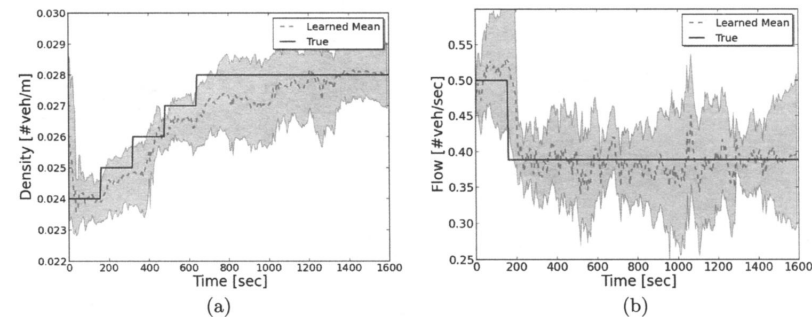


FIG. 13. *True and learned values of parameters and 90% confidence interval band. (a) Critical density learning. (b) Capacity learning.*

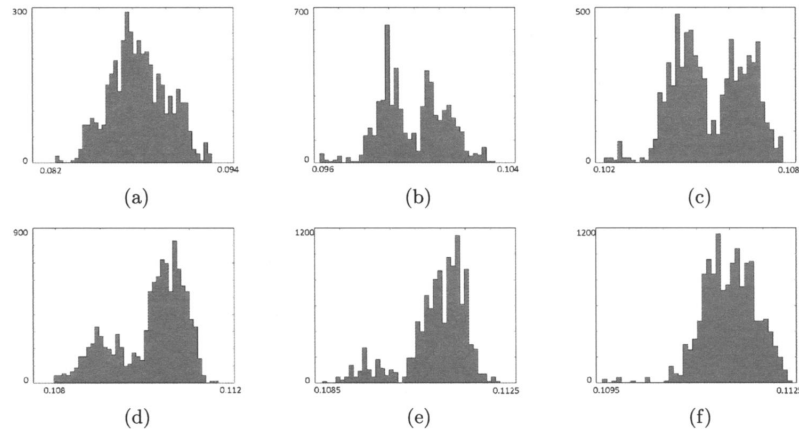


FIG. 14. Uncertainty distribution about traffic flow density estimated by filtering algorithm at cell 3. Panel (a) shows distribution at the time right before the shock wave reaches the cell. Panels (b)–(e) show the time when the shock wave travels through the cell. Panel (f) shows distribution at the time after the shock wave moves beyond the cell. (a) $t = 565$, (b) $t = 615$, (c) $t = 645$, (d) $t = 680$, (e) $t = 690$, (f) $t = 700$.

6. Discussion. In this paper we analyze the LWR traffic flow model with application to Chicago's interstate I-55 highway. We show how particle filtering and learning provides a real-time estimate of the density states. We sequentially learn the parameters of the fundamental diagram which is the central input for the LWR dynamics of traffic flow. Our results have a number of important implications for transportation system management applications. In particular, a real-time assessment of model states and parameters corrects the biases in estimating the current density of states used for forecasting.

Our methodology quickly handles the drop in capacity due to a major traffic accident on Chicago's interstate I-55 highway. We also use a calibration study to show how close a filtered state vector is to the true one. When measurements are sparse in space and the parameters are fixed, pure filtering would misestimate the current state. However, our approach corrects this by incorporating parameters learning simultaneously. This leads to an accurate estimation of traffic density.

There are a number of possible avenues for a future approach. First, the LWR model is only valid when the relationship between flow and density is time independent. Second, the model does not describe traffic behavior within a queue or when particular instabilities such as stop and go traffic exist. Third, the model is not realistic for free-flowing traffic, as vehicle bypassing that happens frequently in this regime is not captured. Although, from a system management perspective, free-flow traffic is not an issue, extending our approach to higher order traffic flow models will lead to improvements in estimation.

Developing methods to incorporate model monitoring is an important area of future research. For example, alternative models might correspond to different assumptions about the shape of the fundamental diagram. We can statistically discriminate two models using a sequential likelihood ratio (Bayes factor), B_t , given by

$$B_t = \frac{p(y_1, \dots, y_t | M_1)}{p(y_1, \dots, y_t | M_0)},$$

where $p(y_1, \dots, y_t | M_i) = \prod_{j=1}^t p(y_j | y_{1:j-1}, M_i)$. This is simply a product of marginal predictive densities, which the particle filter approximates by

$$p^N(y_j | y_{1:j-1}, M) = \frac{1}{N} \sum_{i=1}^N p(y_j | \theta_{j-1}^i, M).$$

Another avenue is to extend our particle algorithm to a transportation network with simultaneous tracking of multiple segments. This will make our methodology applicable to real-life transportation networks of a large metropolitan area. Within our framework, it is feasible to filter over the boundary conditions. This also applies in the case of GPS probes, where inferring the boundary conditions is a hard task, since location and time of the measurement are random and one rarely observes the boundary conditions.

APPENDIX A: DERIVATION OF FLOW MODEL

Let $q(x, t)$, $\rho(x, t)$ and $v(x, t)$ denote traffic flow, density and speed at position x at time t . Kinematic wave theory establishes a relationship between density ρ and flow q , which is known as the *fundamental diagram* given by the functional equation $q(x, t) = q(\rho(x, t))$ where $q(x, t)$ is flow. The conservation law implies that with no inflow or outflow

$$(A.1) \quad \frac{\partial \rho(x, t)}{\partial t} + \frac{\partial q(x, t)}{\partial x} = 0.$$

Combined with fundamental diagram function, we obtain the equation for $\rho(x, t)$:

$$\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial q(\rho)}{\partial \rho} \frac{\partial \rho(x, t)}{\partial x} = 0.$$

The term $w = \partial q(\rho) / \partial \rho$ is called the wave velocity. To get a more intuitive understanding of the problem, it is convenient to use the cumulative flow $N(x, t)$, the number of vehicles that pass location x by time t . Then the conservation law can be derived by evaluating

$$\frac{\partial N}{\partial t} = q(x, t), \quad \frac{\partial N}{\partial x} = -\rho(x, t).$$

Assuming that $N(x, t)$ is smooth,

$$\frac{\partial^2 N}{\partial x \partial t} = \frac{\partial^2 N}{\partial t \partial x},$$

we get the conservation law (A.1). In practice, the function N has discontinuity of the first kind (first derivative), however, the conservation law holds in the case of discontinuities as long as $N(x, t)$ is continuous along the shock path. The method of characteristics can be used to solve the equation (A.1). Specifically, from (A.1) $\rho(x, t)$ is constant ($d\rho/ds = 0$) along a characteristic curve (wave) described by

$$\frac{dt}{ds} = q'(\rho).$$

Eliminating s gives

$$\rho(x, t) = \rho(x - q'(\rho_0)t).$$

Thus, density is constant along the straight line with slope $dq/d\rho$ (characteristic line) and the slope is nothing but a shock propagation speed. For a free-flow speed the shock moves forward and for jammed traffic it moves backward. In Newell's case the forward shock propagation speed is v_f and the backward shock propagation speed is given by w .

APPENDIX B: DERIVATION OF KALMAN RECURSION

Measurements are taken at the first and last cells of the road segment and noise is independently distributed, with covariance structure $V_t = V = vI_2$ and $W_t = W = wI_5$. The operator H_t and the Kalman gain matrix K_t are of the following form:

$$H_t = H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad K_t = K = \begin{pmatrix} \frac{w}{v+w} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{w}{v+w} \\ 0 & 0 \end{pmatrix}.$$

This leads us to the following Kalman updates:

$$C_{t+1} = \begin{pmatrix} w\left(1 - \frac{w}{v+w}\right) & 0 & 0 & 0 & 0 \\ 0 & w & 0 & 0 & 0 \\ 0 & 0 & w & 0 & 0 \\ 0 & 0 & 0 & w & 0 \\ 0 & 0 & 0 & 0 & w\left(1 - \frac{w}{v+w}\right) \end{pmatrix},$$

$$\mu_{t+1} = \left(\mu_1^f + \frac{w(y_1 - \mu_1^f)}{v+w}, \mu_2^f, \mu_3^f, \mu_4^f, \mu_5^f + \frac{w(y_2 - \mu_5^f)}{v+w} \right)^T.$$

The variance of the predictive likelihood distribution is given by

$$HWH^T + V = \begin{pmatrix} v+w & 0 \\ 0 & v+w \end{pmatrix}.$$

REFERENCES

- ANACLETO, O., QUEEN, C. and ALBERS, C. J. (2013). Multivariate forecasting of road traffic flows in the presence of heteroscedasticity and measurement errors. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **62** 251–270. MR3045876
- ARNOTT, R., DE PALMA, A. and LINDSEY, R. (1991). Does providing information to drivers reduce traffic congestion? *Transportation Research Part A: General* **25** 309–318.
- BENGTTSSON, T., BICKEL, P. and LI, B. (2008). Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems. In *Probability and Statistics: Essays in Honor of David A. Freedman. Inst. Math. Stat. Collect.* **2** 316–334. IMS, Beachwood, OH. MR2459957
- BRILON, W., GEISTEFELDT, J. and REGLER, M. (2005). Reliability of freeway traffic flow: A stochastic concept of capacity. In *Proceedings of the 16th International Symposium on Transportation and Traffic Theory* **125143**. College Park, MD.
- CARPENTER, J., CLIFFORD, P. and FEARNHEAD, P. (1999). Improved particle filter for nonlinear problems. *IEE Proc. Radar Sonar Navig.* **146** 2–7.
- CARVALHO, C. M., JOHANNES, M. S., LOPES, H. F. and POLSON, N. G. (2010). Particle learning and smoothing. *Statist. Sci.* **25** 88–106. MR2741816
- CBS CHICAGO (2014). Big delay on interstate 55 after truck fire in Romeoville. Romeoville.
- CHIOU, J.-M. (2012). Dynamical functional prediction and classification, with application to traffic flow prediction. *Ann. Appl. Stat.* **6** 1588–1614. MR3058676
- CHIOU, Y.-C., LAN, L. W. and TSENG, C.-M. (2014). A novel method to predict traffic features based on rolling self-structured traffic patterns. *J. Intell. Transp. Syst.* **18** 352–366.
- CHORUS, C. G., MOLIN, E. J. and VAN WEE, B. (2006). Use and effects of advanced traveller information services (ATIS): A review of the literature. *Transp. Rev.* **26** 127–149.
- CHU, K.-C., YANG, L., SAIGAL, R. and SAITOU, K. (2011). Validation of stochastic traffic flow model with microscopic traffic simulation. In *IEEE Conference on Automation Science and Engineering (CASE)* 672–677. IEEE, New York.
- CLAUDEL, C. G. and BAYEN, A. M. (2010). Lax–Hopf based incorporation of internal boundary conditions into Hamilton–Jacobi equation. Part I: Theory. *IEEE Trans. Automat. Control* **55** 1142–1157. MR2642079
- COCLITE, G. M., GARAVELLO, M. and PICCOLI, B. (2005). Traffic flow on a road network. *SIAM J. Math. Anal.* **36** 1862–1886. MR2178224
- COURANT, R., FRIEDRICH, K. and LEWY, H. (1928). Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.* **100** 32–74. MR1512478
- DAGANZO, C. F. (1995). The cell transmission model, part II: Network traffic. *Transp. Res., Part B: Methodol.* **29** 79–93.
- DERVISOGLU, G., GOMES, G., KWON, J., HOROWITZ, R. and VARAIYA, P. (2009). Automatic calibration of the fundamental diagram and empirical observations on capacity. In *Transportation Research Board 88th Annual Meeting* **15**. Washington, DC.
- DOUCET, A., GODSILL, S. and ANDRIEU, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* **10** 197–208.
- GARAVELLO, M. and PICCOLI, B. (2006). *Traffic Flow on Networks*. American Institute of Mathematical Sciences (AIMS), Springfield, MO. MR2328174
- GAZIS, D. C. and KNAPP, C. H. (1971). On-line estimation of traffic densities from time-series of flow and speed data. *Transp. Sci.* **5** 283–301.

- GODSILL, S. J., DOUCET, A. and WEST, M. (2004). Monte Carlo smoothing for nonlinear times series. *J. Amer. Statist. Assoc.* **99** 156–168. MR2054295
- GODUNOV, S. K. (1959). A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Mat. Sb.* **47 (89)** 271–306. MR0119433
- GORDON, N. J., SALMOND, D. J. and SMITH, A. F. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)* **140** 107–113. IET.
- HOLDEN, H. and RISEBRO, N. H. (1995). A mathematical model of traffic flow on a network of unidirectional roads. *SIAM J. Math. Anal.* **26** 999–1017. MR1338371
- KNOOP, V. L., HOOGENDOORN, S. P. and VAN ZUYLEN, H. J. (2008). Capacity reduction at incidents: Empirical data collected from a helicopter. *Transportation Research Record: Journal of the Transportation Research Board* **2071** 19–25.
- LEBACQUE, J.-P. (2005). First-order macroscopic traffic flow models: Intersection modeling, network modeling. In *Transportation and Traffic Theory. Flow, Dynamics and Human Interaction. 16th International Symposium on Transportation and Traffic Theory*. College Park, MD.
- LEVEQUE, R. J. (2002). *Finite Volume Methods for Hyperbolic Problems*. Cambridge Univ. Press, Cambridge. MR1925043
- LIGHTHILL, M. J. and WHITHAM, G. B. (1955). On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proc. Roy. Soc. London. Ser. A.* **229** 317–345. MR0072606
- LIU, J. and WEST, M. (2001). Combined parameter and state estimation in simulation-based filtering. In *Sequential Monte Carlo Methods in Practice* 197–223. Springer, New York. MR1847793
- MAY, A. D. (1990). *Traffic Flow Fundamentals*. Prentice Hall, New York.
- MIHAYLOVA, L., BOEL, R. and HEGYI, A. (2007). Freeway traffic estimation within particle filtering framework. *Automatica J. IFAC* **43** 290–300. MR2281833
- MURALIDHARAN, A. and HOROWITZ, R. (2009). Imputation of ramp flow data for freeway traffic simulation. *Transportation Research Record: Journal of the Transportation Research Board* **2099** 58–64.
- PITT, M. K. and SHEPHARD, N. (1999). Filtering via simulation: Auxiliary particle filters. *J. Amer. Statist. Assoc.* **94** 590–599. MR1702328
- RICHARDS, P. I. (1956). Shock waves on the highway. *Operations Res.* **4** 42–51. MR0075522
- SCHREITER, T., VAN HINSBERGEN, C., ZUURBIER, F., VAN LINT, H. and HOOGENDOORN, S. (2010). Data-model synchronization in extended Kalman filters for accurate online traffic state estimation. In *Proceedings of the Traffic Flow Theory Conference* **86**. Annecy, France.
- SNYDER, C. (2011). Particle filters, the “optimal” proposal and high-dimensional systems. In *Proceedings of the ECMWF Seminar on Data Assimilation for Atmosphere and Ocean*. Shinfield Park, Reading.
- STORVIK, G. (2002). Particle filters for state-space models with the presence of unknown static parameters. *IEEE Trans. Signal Process.* **50** 281–289.
- SUN, X., MUÑOZ, L. and HOROWITZ, R. (2003). Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control. In *Proceedings of 42nd IEEE Conference on Decision and Control* **6** 6333–6338. IEEE, New York.
- TEBALDI, C. and WEST, M. (1998). Bayesian inference on network traffic using link count data. *J. Amer. Statist. Assoc.* **93** 557–576. MR1631325
- TRANSPORTATION RESEARCH BOARD (2010). Highway capacity manual. National Academies of Sciences, Engineering, and Medicine, Washington, DC.
- WANG, Y. and PAPAGEORGIOU, M. (2005). Real-time freeway traffic state estimation based on extended Kalman filter: A general approach. *Transp. Res., Part B: Methodol.* **39** 141–167.
- WESTGATE, B. S., WOODARD, D. B., MATTESON, D. S. and HENDERSON, S. G. (2013). Travel time estimation for ambulances using Bayesian data augmentation. *Ann. Appl. Stat.* **7** 1139–1161. MR3113504

WORK, D. B., TOSSAVAINEN, O.-P., BLANDIN, S., BAYEN, A. M., IWUCHUKWU, T. and TRAC-
TON, K. (2008). An ensemble Kalman filtering approach to highway traffic estimation using GPS
enabled mobile devices. In *CDC 2008 of 47th IEEE Conference on Decision and Control* 5062–
5068. IEEE, New York.

BOOTH SCHOOL OF BUSINESS
UNIVERSITY OF CHICAGO
CHICAGO, ILLINOIS 60637
USA
E-MAIL: ngp@chicagobooth.edu

ARGONNE NATIONAL LABORATORY
LEMONT, ILLINOIS 60439
USA
E-MAIL: vs@anl.gov