ARTICLE IN PRESS

International Journal of Transportation Science and Technology xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Transportation Science and Technology

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

Cycle-to-cycle queue length estimation from connected vehicles with filtering on primary parameters

Gurcan Comert [a,b,*], Negash Begashaw [a]

[a] Computer Science, Physics, and Engineering Department, Benedict College, Columbia, SC 29204, USA
[b] Information Trust Institute, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

A R T I C L E I N F O

Article history:
Received 18 November 2020
Received in revised form 10 February 2021
Accepted 19 April 2021
Available online xxxx

Keywords:
Queue length estimation
Kalman filter
Particle filter
Connected vehicles
Market penetration rate

A B S T R A C T

Estimation models from connected vehicles often assume low level parameters such as arrival rates and market penetration rates as known or estimate them in real-time. At low market penetration rates, such parameter estimators produce large errors making estimated queue lengths inefficient for control or operations applications. In order to improve accuracy of low level parameter estimations, this study investigates the impact of connected vehicles information filtering on queue length estimation models. Filters are used as multilevel real-time estimators. Accuracy is tested against known arrival rate and market penetration rate scenarios using microsimulations. To understand the effectiveness for short-term or for dynamic processes, arrival rates, and market penetration rates are changed every 15 min. The results show that with Kalman and Particle filters, parameter estimators are able to find the true values within 15 min and meet and surpass the accuracy of known parameter scenarios especially for low market penetration rates. In addition, using last known estimated queue lengths when no connected vehicle is present performs better than inputting average estimated values. Moreover, the study shows that both filtering algorithms are suitable for real-time applications that require less than 0.1 second computational time.

1. Introduction

Queue length estimation (QLE) with sensor technology has been getting more attention as real-time estimation is becoming feasible. Queue lengths can be utilized especially in intelligent transportation system operations and control. In this usage of queue lengths, market penetration levels, coverage of sensors, and accuracy are crucial points to consider. Given low penetration levels, researchers aim to improve estimators by filtering (Tiaprasert et al., 2015; Yin et al., 2018; Emami et al., 2019) as well as fusing with existing surveillance technologies (e.g., inductive loop, video, piezoelectric, and range sensors). Real-time performance and simplicity would facilitate methods to be used in simpler controllers (Khan et al., 2017). Change in low level parameters is also a concern as known and constant parameter assumptions would generate inaccurate estimations at low market penetration rates (MPR) or coverage of surveillance systems (e.g., connected vehicles, inductive loop, or video cameras). Although, fusion and data driven methods would be alternative to improve estimators under low coverage (Li et al., 2013; Grumert and Tapani, 2018; Aljamal et al., 2020), we focused to single type of sensing technology

* Corresponding author.
E-mail address: gurcan.comert@benedict.edu (G. Comert).

https://doi.org/10.1016/j.ijtst.2021.04.009
2046-0430/

Please cite this article as: G. Comert and N. Begashaw, Cycle-to-cycle queue length estimation from connected vehicles with filtering on primary parameters, International Journal of Transportation Science and Technology, https://doi.org/10.1016/j.ijtst.2021.04.009

(i.e., connected vehicles). Researchers have used filtering to reduce error in such cases (Yin et al., 2018; Emami et al., 2019). Filtering can be applied in different layers depending on the objective. Since true performance measures (e.g., queue lengths) are not known, assuming relatively short unchanging parameters would help to improve estimators. The challenges are multimodality over longer term and underlying unknown distribution for parameters. In this paper, we utilize Kalman and Sequential Monte Carlo (e.g., bootstrap, particle) filters to estimate primary parameters more accurately and thereby improve cycle-to-cycle queue length estimation at traffic intersections within the framework of connected vehicles (CV).

Readers are referred to the survey papers by Zhao et al. (2019) and Guo et al. (2019) for general queue length estimation from connected vehicles at traffic intersections, and Asanjarani et al. (2017) for queue length estimation. Recent similar studies include methods reconstructing trajectories (Xu et al., 2017; Rompis et al., 2018; Zhao et al., 2019), cycle-to-cycle estimations (Tan et al., 2019), and real-time applications with high market penetration rates (Mei et al., 2019). Based on shockwaves and queue dynamics, Yin et al. showed effectiveness of Kalman filter in QLE, especially within 20% MPR (Yin et al., 2018). Discrete Fourier transform based filtering was applied by Tiaprasert et al. (2015) on queue length estimation but not on primary parameters. Any overestimation and underestimation on primary parameters would be passed to QLE accuracy. Queue sizes, especially medium to high volume-to-capacity ratios can fluctuate significantly. Filtering estimation QLs would result as low or high trimmed values. Thus, knowing relatively fixed parameters at a level can be better target to filter.

Stochastic models of QLEs from CVs mainly concerns with stationary (behavior after sufficiently long time interval) or non-stationary (short-time intervals, by cycle, and red phase). Such classification is impacted by treatment of arrival rate, distribution, and market penetration rate (MPR) (probe percentage). For instance, Yang and Menendez showed that 20% to 30% MPR is required to control signals (Yang and Menendez, 2018). Although 20 to 30% MPR is shown to be required to control signals (Yang and Menendez, 2018), in another study Argoteet al. showed different accuracy levels, $\pm 10\%$ with 80% MPR (Argote et al., 2011). Using sensor fusion with the existing deployed traffic sensing technology (*e.g.*, inductive loops, license plate recognition) was also investigated by the researchers (Wu and Yang, 2013; Li et al., 2013; Comert, 2013; Tan et al., 2020; Mandal et al., 2020).

Even if penetration levels increase, efficient estimators performing at low MPRs would still be desirable for efficient data management. The question of how much data one needs is important not to overload systems for safety and security reasons. Partially supporting these, in an interesting report (Arizona et al., 2016), researchers have found in a simulation environment that over 200 vehicles at 100% penetration level signal control from CVs becomes difficult due to computational overhead. Rural roads and isolated intersections would also have lower MPRs. Regardless, increasing QLEs and filtering inconsistent extreme over and under estimations are very important for traffic operations and any partially observed systems.

1.1. Contributions of this study

Simple expectation-based QLE models and estimators with known characteristics in Comert (2016) enable application of filtering algorithms to deal with measurement of low MPR related errors. The estimation problem where we only assume partially observed system becomes interesting. This approach can be adopted for any similar problem like autonomous vehicles. Studies in estimation of primary parameters using only fundamental CV information (*i.e.*, time, location, speed, and vehicle type etc.) are getting more attention (Comert, 2016; Van and Farhi, 2018; Zhao et al., 2019). But, there are limitations for connected vehicles with low MPRs.

In this study, we aim to improve accuracy by enhancing the low level parameter estimators using filtering algorithms. The resulting formulae would not depend on any primary parameters such as arrival rate or MPR. So, the method can be used for time dependent arrival rates, MPR, or under dynamic conditions.

2. Problem definition

Estimation of queue lengths is one of the inherent key capabilities of connected and autonomous vehicles for operations control as moving (as opposed to fixed loops or video cameras) objects reporting sample observations from traffic systems. In this study, we aim to estimate the queue lengths at traffic intersections without knowing the true low level parameters as well as true queue lengths through classic localization idea in robotics (Thrun, 2002; Thrun et al., 2006). Assume that CVs can share location and time information with a time reference point (e.g., signal timing, red duration), number of queued vehicles is aimed.

Considering Fig. 1 as an example layout, for any approach right before the signal turns green, we aim to estimate the queue length. For instance, consider left turn on east–west bound approach. Assuming that the blue or darker cars are connected vehicles, we try to find total queue length of $N = 7$ given last connected vehicle's location $L = 6$, number of connected vehicles $M = 2$, and arrival time during red $T = 10$ seconds s with remaining $R - T$ s.

Mathematically, the estimator $E(N_i|L = l_i, T = t_i, M = m_i, R)$ for short target i time intervals is aimed to be described. In the estimator, N_i is cycle-to-cycle total queue lengths at the end of analysis period, L_i location of the last CV, T_i is queue joining time of the CV, and M_i is number of CVs in the queue. Using models in the literature, for Poisson arrivals, simple $l_i + (1 - p)\lambda(R - t_i)$ are shown to be an unbiased estimators. Certainly, arrival rate λ and market penetration rate p are equally difficult to determine in practice. So, several estimators for λ and p are also given in Comert (2016). However, for low market penetration levels less than 30%, estimators tend to perform with high variance. Thus, it is intuitive to utilize filters similar to

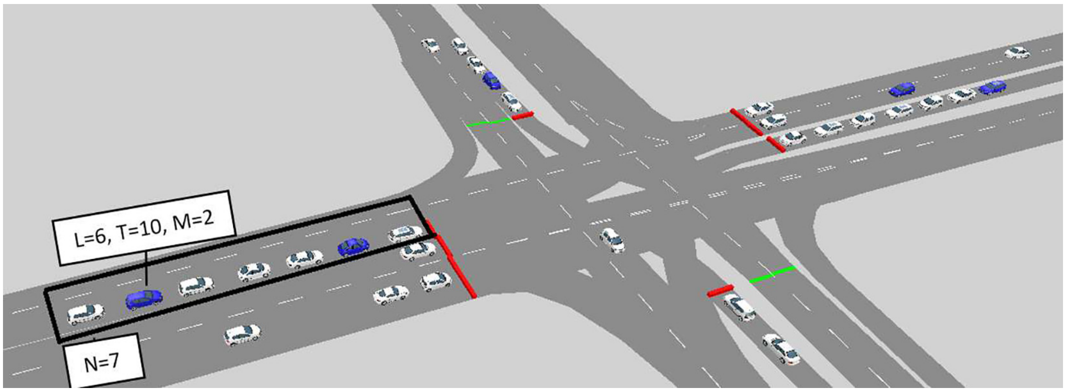

Fig. 1. An example intersection with multiple approaches generated in Vissim.

localization idea in robotics. The difference in this problem is that there are no landmarks assumed to correct measurements. It is advantageous to assume constant market penetration and arrival rate for relatively short terms, e.g., 15 to 30 min. This assumption would allow us to locate true values from unimodal distributions. It can be relaxed using desirably nonparametric filtering techniques that would allow multimodal distributions. The nature of the problem can be summarized as follows:

i. Queue lengths show nonlinear behavior with unknown distribution type.
ii. The underlying process model is also very complex depending on arrival rate, changing signal timing, day of time, and roadway geometry etc.

Alternative to such conditions can be listed as discrete Bayes filters and particle filters. Since we are working with constant parameters within short time intervals such as 15 min, we aim to filter parameters for unimodal distribution case. If we are able to get $\left[\hat{p}, \hat{\lambda}\right]$ values as correctly as possible, then we would be able to estimate QLs more correctly and do not violate the constant parameter within short intervals. In the figures presented in Section 4, estimated parameters are given in incremental time interval.

3. Methodology

In this paper, filters are mainly used to estimate primary parameters, arrival rate (λ) and market penetration level (p). Impact of filtering on the lowest level CV time, location, and count information was also checked. Parameter filters would perform better under given consistent behavior of noise levels and described underlying dynamics such as object movement. In this case, no assumption was made for underlying distribution especially for particle filter as it is considered a distribution-free multimodal method. Whereas, KF assumes Gaussian and unimodal underlying process distribution. For both methods, queue length estimation scenarios would be appropriate since unimodal/constant higher level arrival rate and market penetration rates are assumed within 15-min time intervals.

In order to set up the application, first consider the estimators in general form of $E(N = n|L = l, T = t, M = m) = l + (1 - p)\lambda(R - t)$. The problem with this estimator is lack of accuracy on $\hat{p}$ and $\hat{\lambda}$, especially at low p ($p \leqslant 0.30$). Note that fusion with other sensing technologies would also be solution, however, it is beyond the scope of this study (Comert, 2013; Comert and Cetin, 2020). We aim to filter and find out true parameter values given short-time intervals *i.e.*, 15 min. Estimator combinations that best performed in QLE (Comert, 2016) were

$$\hat{p} = \left\{\left(1 - \frac{m_i t_i}{m_i t_i + (l_i - m_i)R}\right), \frac{m_i}{l_i}\right\}$$
$$\hat{\lambda} = \left\{\left(\frac{l_i - m_i}{t_i} + \frac{m_i}{R}\right), \frac{l_i}{t_i}\right\}$$

Thus, queue length estimators are used in Eqs. (2) and (3) for comparison of simple estimator in Eq. (1) with $\hat{p} = \frac{m_i}{(l_i/t_i)R}$ and $\hat{\lambda} = l_i/t_i$. Note that this estimator simplifies to $\frac{Rl_i}{t_i} - m_i(R - t_i) = R\hat{\lambda} - m_i(R - t_i)$.

$$E\left(\hat{N}_i|l_i, t_i, m_i, q_i\right) = \begin{cases} l_i + \left(1 - \frac{m_i t_i}{l_i R}\right)\left(\frac{i}{t_i}\right)(R - t_i), & l_i > 0 \\ E\left(\hat{N}_j|l_j, t_j, m_j, q_j\right) \neq 0, & j = 1, \cdots, (i-1), l_i = 0 \end{cases} \tag{1}$$

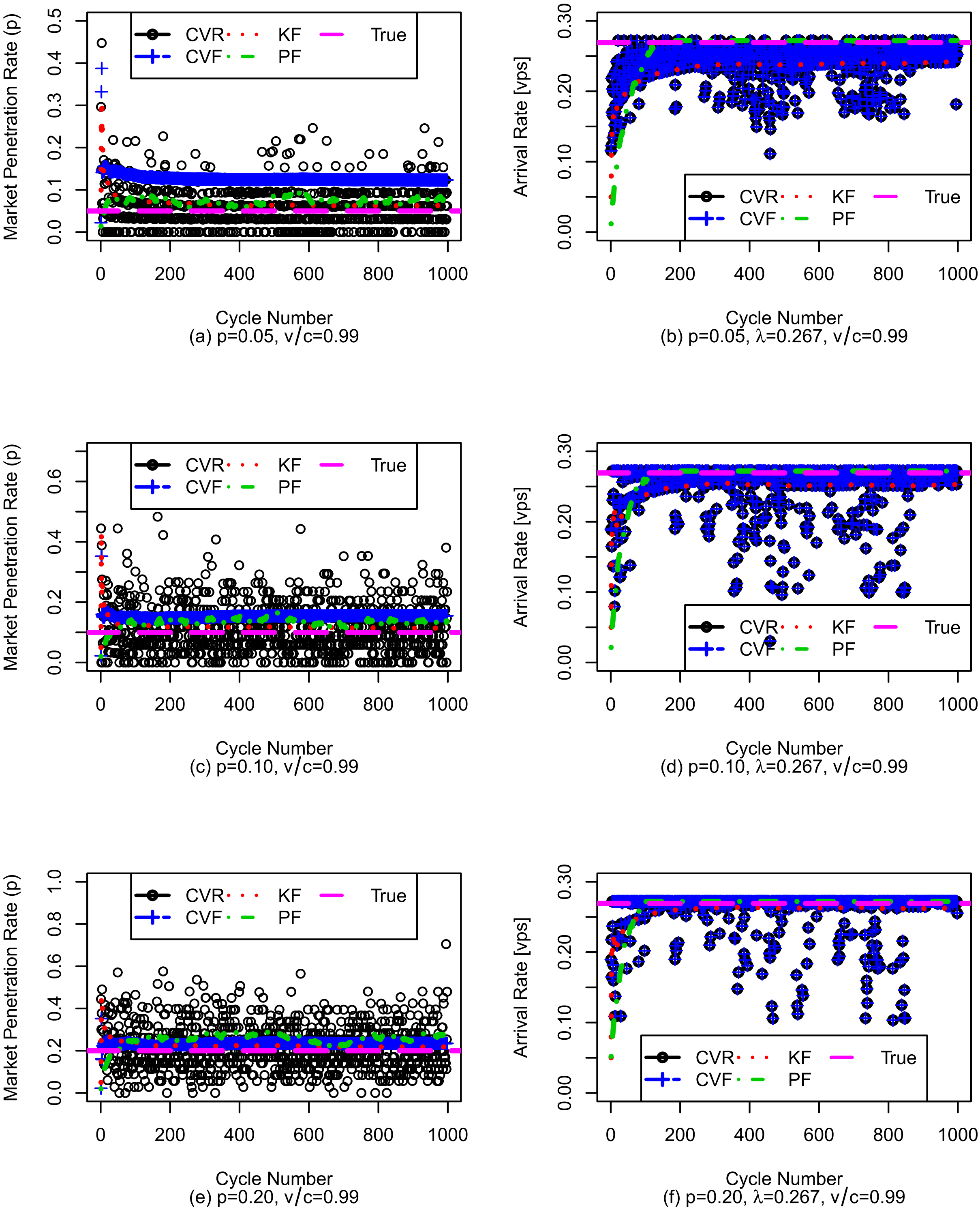

Fig. 2. Parameter estimation example with filters at p, $v/c = 0.99$ as average of 3 random seeds.

$$E\left(\hat{N}_i | l_i, t_i, m_i, q_i\right) = l + \left(1 - \frac{m_i}{l_i}\right)\left(\frac{l_i}{t_i}\right)(R - t) = l_i + \left(\frac{l_i - m_i}{t_i}\right)(R - t_i) \tag{2}$$

$$E\left(\hat{N}_i | l_i, t_i, m_i, q_i\right) = l + \left(1 - \frac{m_i t_i}{m_i t_i + (l_i - m_i)R}\right)\left(\frac{l_i - m_i}{t_i} + \frac{m_i}{R}\right)(R - t_i) = m_i + \frac{R(l_i - m_i)}{t_i} \tag{3}$$

3.1. Kalman filter

In its simplest form, Kalman filter (KF) assumes underlying Gaussian measurements or noise (Labbe, 2014). KF also performs well if a physical system is modeled adequately (e.g., vehicle moving dynamics). In the QLE problem, both are not true. We may have different parameter values and no closed-form formulas are available for dynamics of parameters (*i.e.*, p, λ) given connected vehicle information. However, it is possible to demonstrate the problem better and can easily be implemented. After careful experimentation, observed CV information for each $\hat{\lambda}$ are defined as $X_i = \min\left\{\left(\frac{L_i - M_i}{T_i} + \frac{M_i}{R}\right), 0.272\right\}$. When lower level L, T are also filtered, in estimators XC_i used as $XC_i = \min\left\{\frac{XL_i}{XT_i}, 0.280\right\}$. They were filtered for all λ, p.

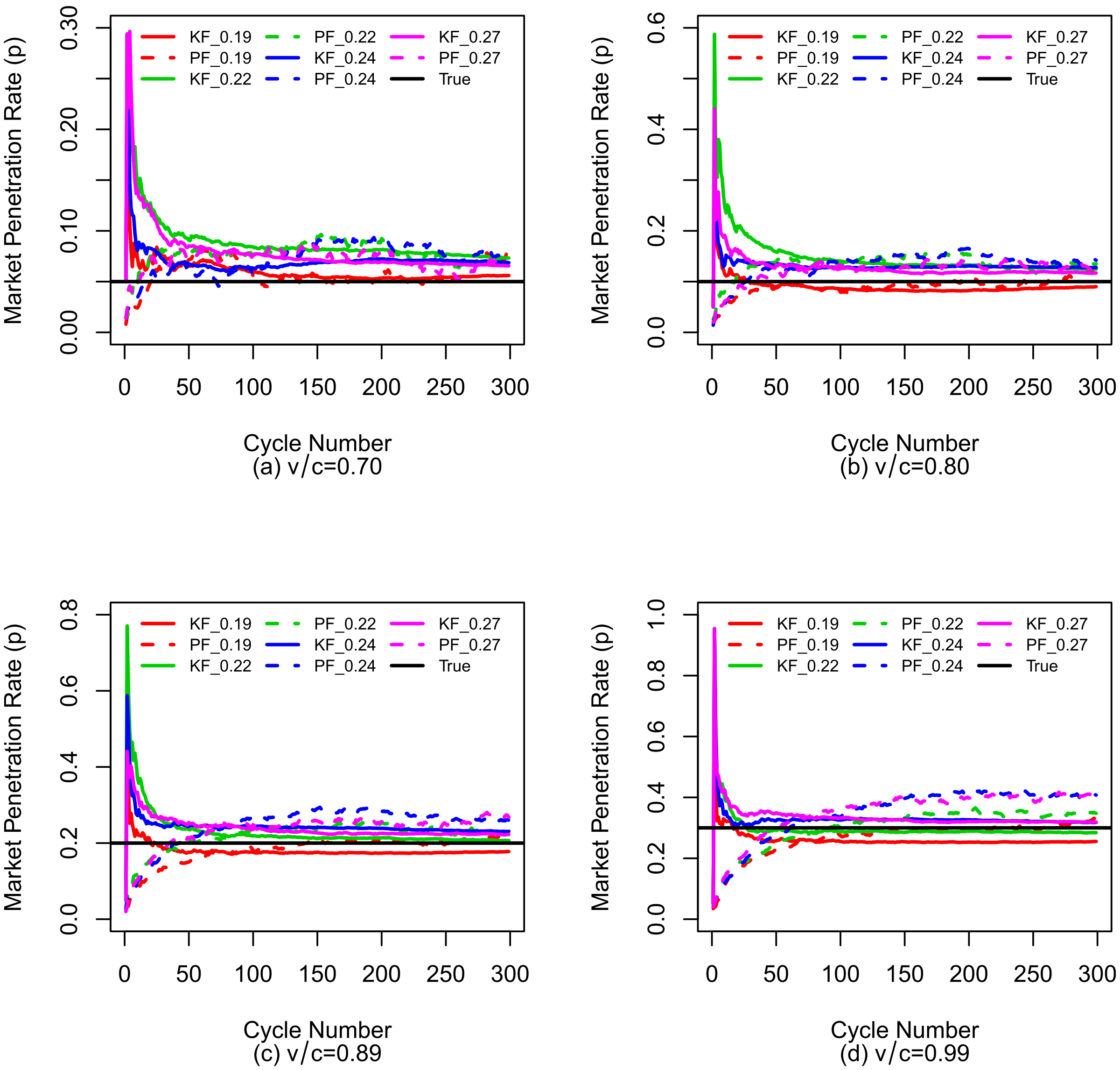

Fig. 3. Estimation examples for p with filters at $p = \{0.05, 0.10, 0.20, 0.30\}$, $v/c = \{0.70, 0.80, 0.89, 0.99\}$ as average of 3 random seeds.

Similarly, market penetration rate ($\hat{p}$) was estimated and filtered using $X_i = \min\left\{\left(1 - \frac{M_i T_i}{M_i T_i + (L_i - M_i)R}\right), 1.00\right\}$ and $XC_i = \min\left\{\frac{XM_i}{XL_i}, 1.00\right\}$.

Algorithm 1: Kalman Filter

assign:
$\hat{X}_i = \mu_i$
$S_{X_i} = S_{X_{i-1}}$
update:
$K_{X_i} = S_{X_i}/(S_{X_i} + s)$
$Y_{X_i} = X_i - \hat{X}_i$
$\mu_{i+1} = \mu_i + K_X Y_X$
$S_{X_{i+1}} = (1 - K_{X_i})S_{X_i}$

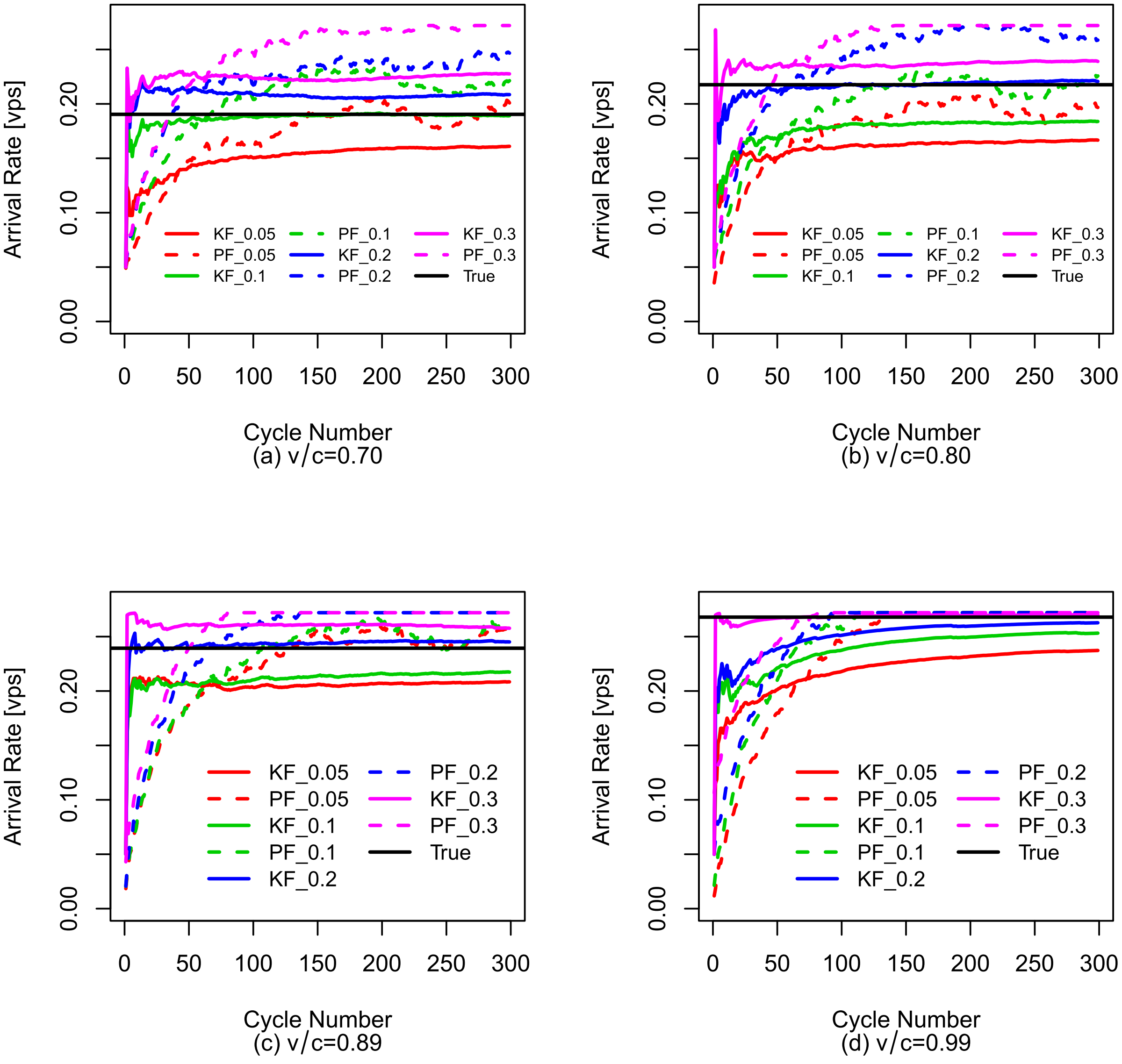

Fig. 4. λ estimation examples with filters at $p = \{0.05, 0.10, 0.20, 0.30\}$, v/c=$\{0.70, 0.80, 0.89, 0.99\}$ as average of 3 random seeds.

In the algorithm, S_{X_i} is the system uncertainty, s is sensor uncertainty, K_{X_i} is the Kalman gain in cycle i, X_i observation in cycle i, Y_{X_i} estimation error, and μ_{i+1} is the state update using Kalman gain and estimation errors.

3.2. Particle filters

Sequential Monte Carlo (SMC) framework contains many similar filters such as particle, bootstrap, and sequential importance filters with slight changes in the resampling (i.e., update) methods (Liu and Chen, 1998; Gustafsson et al., 2002; Doucet and Johansen, 2009; Carvalho et al., 2010; Lopes and Tsay, 2011; Bernardo et al., 2011). QLE setting has $[p, \lambda]$ with corresponding weights each showing how likely they represent true $\hat{p}$ and $\hat{\lambda}$ values. Unlike usual robotics localization problem landmarks, sensing, or true (or noisy) parameter values are not available in QLE application for low MPRs. Thus, after each movement, convergence to true parameters (importance resampling) is more challenging. Heuristic SMC filtering steps can be written within parameter estimation concepts as follows:

(1) Observe new information from the cycle as triplets $[M, L, T]$ and estimate with $[\hat{\lambda}, \hat{p}]$.
(2) Update the weights of the particles based on the new estimates from the updated observations.

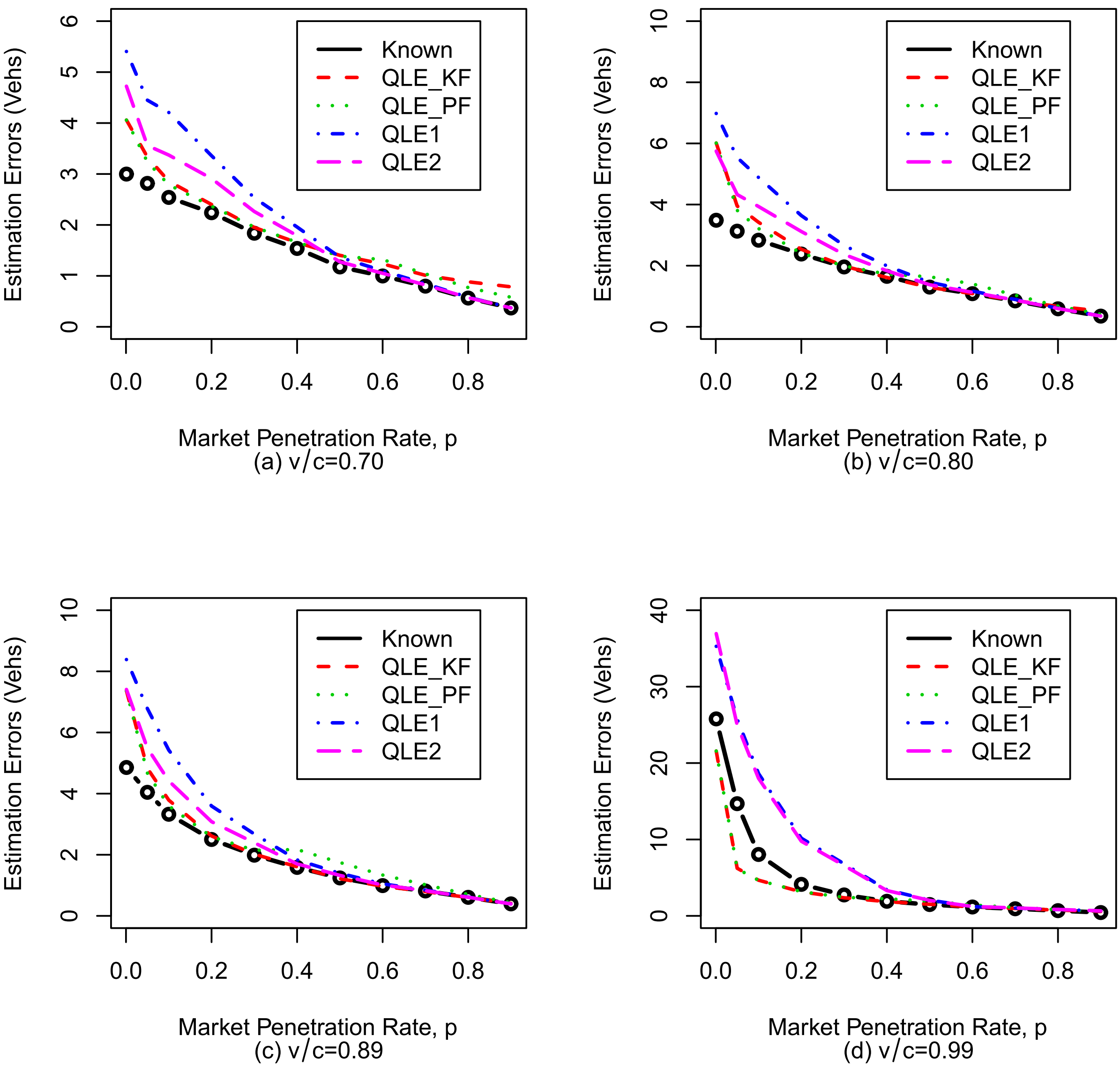

Fig. 5. Performance with filters for different p, λ, and v/c.

(3) Based on the triplet received and new estimation of parameters, resample with replacement to eliminate particles that are highly unlikely. Particles close to the estimates get higher weights.
(4) Estimate weighted mean and variance of the filtered $\hat{p}, \hat{\lambda}$.

Algorithm 2 is adopted from Lopes (2009). Initial weights are input as uniform, observations are denoted by $X_{i+1} \sim p(X_{i+1}|Z_{i+1}, \lambda)$, and transitions are denoted by $X_{i+1} \sim p(X_{i+1}|X_i, \lambda)$. Posterior distributions with $p\left(Z_i, \lambda|X^i\right)$ with $X^i = (X_1, \ldots, X_i)$. Since, we do not have movement for particles, a particle filter (Bootstrap filter) with a simpler resampling wheel can be written as propagate: $\left\{Z_i^{(j)}\right\}_{j=1}^N$ to $\left\{\widetilde{Z}_{i+1}^{(j)}\right\}_{j=1}^N$ via $p\left(Z_{i+1}|Z^i\right)$, generate: $\left\{Z_{i+1}^{(j)}\right\}_{j=1}^N$ to $\left\{\widetilde{Z}_{i+1}^{(j)}\right\}_{j=1}^N$ via $p\left(Z_{i+1}|Z^i\right)$, and resample with weights: $w_{i+1}^{(j)} \propto p\left(X_{i+1}|\widetilde{Z}_{i+1}^{(j)}\right)$. In fact, without approximately defined dynamics of particles, full version of particle filter performed very similar with higher computational run time.

Algorithm 2: Particle Filter

N particles: $X \sim \mathcal{N}(0.0, 0.05)$, $\widehat{X}_{cycles \times n}$
for *cycle i* **do**
particle jittering: $X\prime = X + \epsilon$
$w = f(X_i, X\prime/(1 + X\prime^2, 1.00))$
sample: $n\ from : X\prime\ with\ p(w)$
end for
$X = X\prime$, $\widehat{X}_{i,1:n} = X$, **median**$(\widehat{X})$

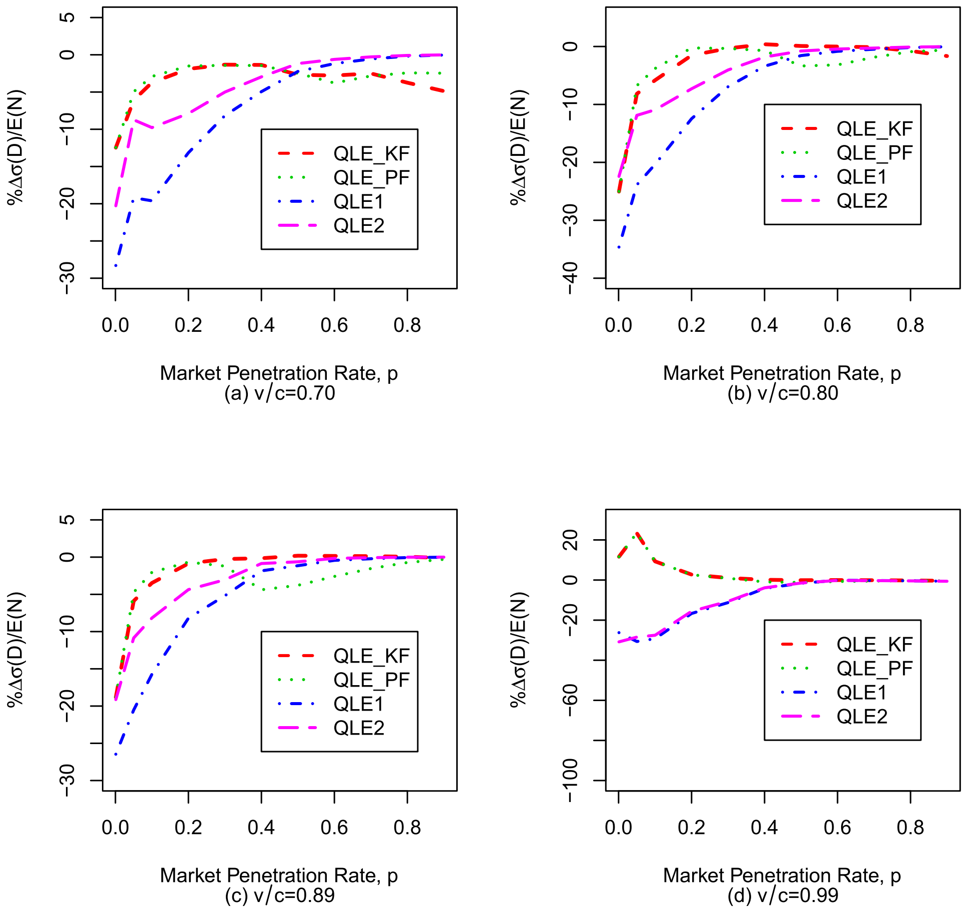

Fig. 6. Impact of filters in estimation $\%\Delta\sqrt{V}(D)/E(N)$ for different p, λ, and v/c average of 3 random seeds.

In Algorithm 2, numerical values are selected after trial and errors on one series (e.g., volume to capacity ratio $v/c = 0.88$ and $p = 0.10$). Low jittering noise was selected to prevent low resampling or fast bad convergence $\epsilon = \mathcal{N}(0.0, 0.0005)$.

4. Numerical Results

Numerical examples are given to show performance of the filtering approach. Analysis include discussions for both filters and queue length estimation from connected vehicles for fixed and varying primary parameters. Simulated intersection queue data are used in the examples. An isolated signal with fixed signal control in Vissim is used to test the estimators with filtered vehicle information after revising the network generated in Comert (2016) and default *queue* definition in Vissim is used (Liu et al., 2019). Simple intersection is used only for generating queueing data. As mentioned earlier estimators can be adopted for approaches with a standing queue.

Note that given required information, estimators can be used under any condition and signal control. Queue lengths that we observe show an autocorrelated times series as dynamic stochastic process. Thus, if estimators perform well here, they can be used under any conditions as only fundamental vehicle information location, time, count, and time period of estimation (e.g., signal timing such as red, green duration) was assumed to be known.

Vehicle trace files generated from microsimulations are used to evaluate queue length estimation without and with initial queue cases for arrival rates of $\lambda = \{0.163, 0.190, 0.218, 0.239, 0.267\}$ vehicle per second (*vps*) that correspond to volume-to-capacity ratios of $v/c = \{0.60, 0.70, 0.80, 0.89, 0.99\}$ and $p = \{0.001, 0.05, 0.10, 0.20, \ldots, 0.90\}$ for the desired simple intersection with 90 s (s) cycle length equal 45 s red/green split with no yellow or all red phases. Note that our approach is QLE given partial information about the system and can be applied for any approach with a standing queue at a certain time period (*e.g.*, red duration).

4.1. Fixed parameters

Cycle-to-cycle updated parameter estimators using filtering that do not involve cumulative connected vehicles information are given in Figs. 2–8. In order to check the impact in lowest level information (*i.e.*, applied on L, T,and M), estimator values with raw CV are denoted by *CVR* and the values with filtered CV information are denoted by *CVF*. Impact is more seen in $\hat{p}$. However, resulting QLEs are not much impacted when using raw or filtered CV values.

From Fig. 2, it can be seen that highly dispersed ratios are filtered to point desired parameter level. However, we are not able to obtain true parameters for $p < 0.20$. At MPR $p = 0.20$, parameter estimators are improved and with filtering, we are able to track the underlying true parameter values p and λ very closely.

Estimator performances for $\hat{\lambda}$ for different p levels and $\hat{p}$ for different λ levels are given in Figs. 3 and 4, respectively. In Fig. 4, KF is able to get closer to true values at low volume-to-capacity ratio $v/c = 0.70$ and $p = 0.10$. In all cases, as p gets higher, estimator performs better. Similarly, in Fig. 3, it can be seen that estimation with filters perform well for all λ and p values. Both KF and PF show similar behavior.

In Fig. 5, estimation errors with and without filtering are presented. Filtering improves the estimation errors compared to unknown primary estimators (denoted by QLE1 and QLE2). Known denotes the estimation with set λ and p values in the simulation runs. With filtering, at 20% market penetration level, errors with known primary parameters can be met. As

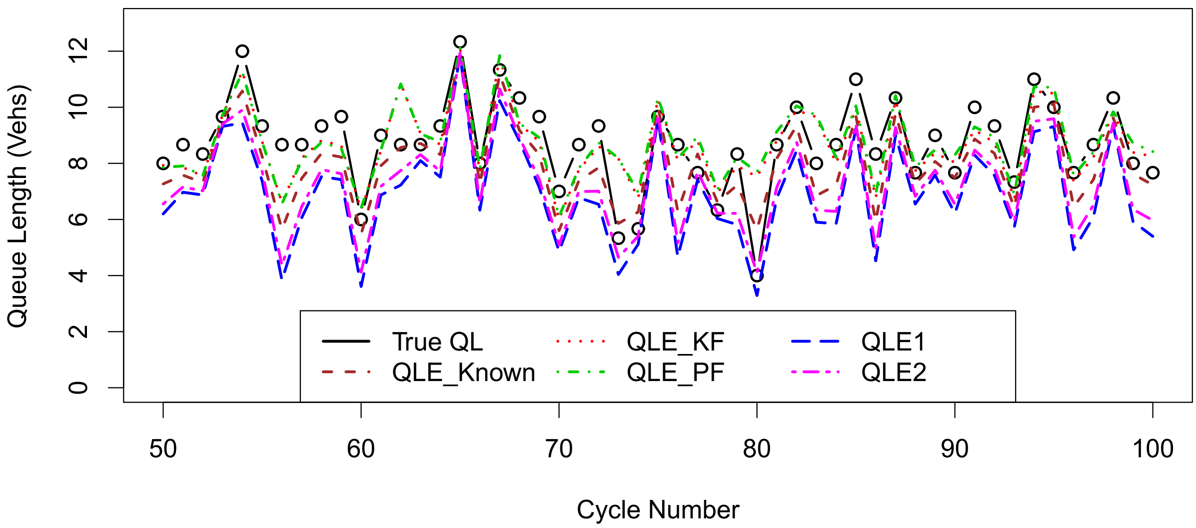

Fig. 7. Estimation example with filters at $p = 5\%$, $v/c = 0.99$ as average of 3 random seeds.

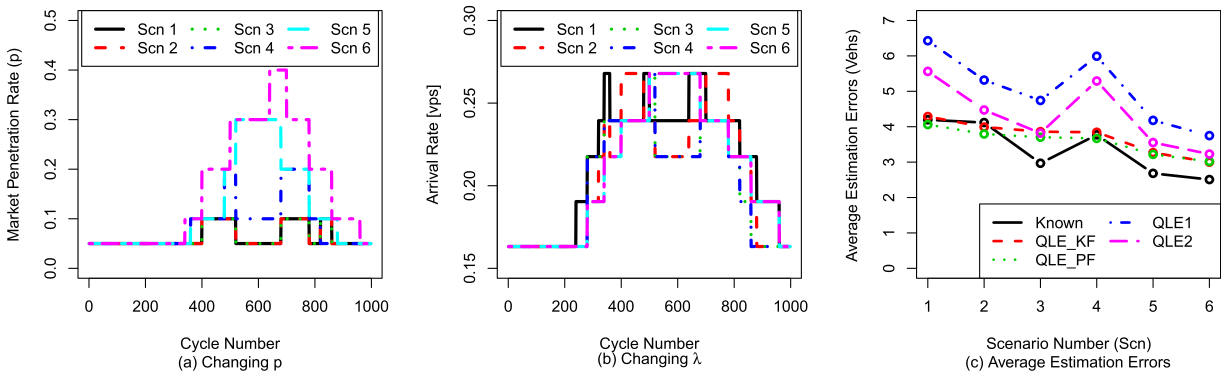

Fig. 8. Scenarios with changing p, λ, and resulting estimation errors in $\sqrt{V(D)}$ [vehs].

volume-to-capacity ratio increases, filtering is able to meet and improve the queue length estimation compared to known case.

The differences in the form of $\%\Delta\sqrt{V}(D)/E(N)$ from QLE with known (p, λ) errors can be seen in Figs. 6a-d. In all v/c values up to 40% MPR, QLE with filtering shows improved errors, *i.e.*, lower than QLE1 and QLE2. This is desirable as accuracy improvement is needed at low MPR values. Only for $v/c = 0.70$, after 50% MPR filtering shows worse errors. However, these errors are already very low (less than 1 vehicle in Fig. 5). QLE with filtering shows better performance than QLE with known (p, λ) up to 40% in $v/c = 0.99$.

Example of queue length estimation is shown in Fig. 7 for $p = 5\%$ and $v/c = 0.99$. Filtered estimators are able to closely follow true QLs. Extreme reactions due to misrepresenting CAVs information are filtered. True queue lengths were obtained by scanning queued vehicles in Vissim at the end of each red phase.

All the examples so far have been on single parameter for relatively long time. Estimators are able to perform well in 50 cycles or 1.25 hours. Next, we consider changing parameter cases and performance of estimation with filters.

4.2. Dynamic parameters

In this subsection, estimators' performance are shown when parameters change for shorter time intervals. Since the physics of queue evolution is not defined in closed form for KF, it is expected that PF performs better for multi-regime parameters as shown in Fig. 8a-b. These scenarios are arbitrarily generated by increasing and decreasing with a peak to mimic changing traffic conditions. Setting up full factor experiments is beyond the scope of this paper.

Overall average estimation errors from these scenarios are also given in Fig. 8c. From the figure, we can deduce that PF performs slightly better than KF in all scenarios. Estimation with filtering outperforms unknown parameter cases and closely follows the estimation errors of known (p, λ) QLE.

Example queue length estimations are given in Fig. 9 for 4 scenarios where overall we can see filtering helps to more closely follow true queue lengths. Behavior of KF and PF on changing parameter estimations are shown in Fig. 10. PF is expected to perform better under dynamic p and λ behaviors. It is also observed that for low p values, using last estimated queue length performs much better than using average values until then. This is due to volatile behavior of queue lengths instead of a level we are able to catch high and low points better with last known point missing data amputation.

For computational times, for a single scenario for 1000 cycles running KF for λ and p estimations on a PC with 8 GB of memory, Pentium I5 Quad-Core CPU took 0.03 s and 0.04 s, respectively. For running PF took 3.50 s and 3.31 s, respectively. As per cycle calculations are well below 0.1 s, both filtering algorithms can be utilized in real-time cycle-to-cycle control and safety critical applications.

4.3. Results from field test data

Effectiveness of the approach is also shown using 2014 ITS World Congress Connected Vehicle Test Bed Demonstration Data (Dataset, 2014). The queue length part of the dataset was collected at the intersection of Larned and Shelby streets, Detroit, Michigan. Between September 8–10, 2014, the data contains 98, 254, and 135 manually collected maximum queue and number of connected vehicles observations where rear of connected vehicles were marked with blue X. Each row of dataset includes hour, minute, second of the observation, the true maximum queue lengths, and the number of connected vehicles in the maximum queues (i.e., M) on left, center, and right lane of Larned street approach.

The dataset does not provide the detailed information that were used (i.e., L, T, M, and R) in the models proposed so far, however, with some assumptions, it is interesting to investigate if the accuracy of queue length estimates can be improved? Therefore, back of queue observations were assumed to be collected at the end of varying red phases. Time between two observations were assumed to be cycle lengths and since this is a major intersection half of cycle lengths were assumed

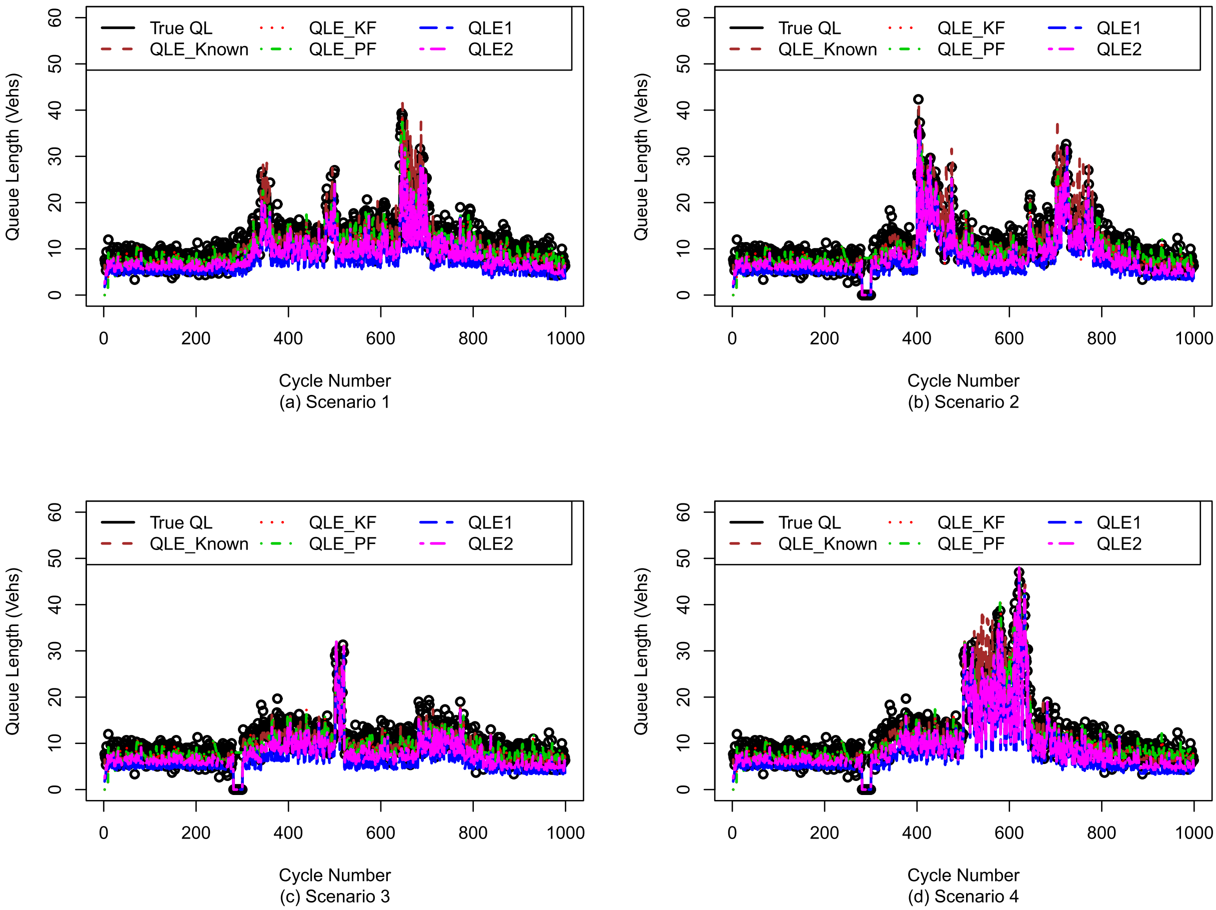

Fig. 9. Queue length estimation examples.

to be red phases. There was no steady growth of queue lengths and many zero values, we assume no initial queues were present with low to medium v/c of 0.50. Capacity of the approach was also approximated by observed maximum queue value of 10 vehicles within 70 s ($10 \times 3600/35$=1029 vehicles per hour or 0.286 vehicle per second (vps) saturation flow rate). These values are used essentially in Highway Capacity Manual (HCM) and back of queue calculations. Note that values would not be reflecting actual capacity and phase splits, however, comparison against true values would provide insights about effectiveness of our approach in addition to the simulation analysis.

The approach was compared with the approximations from HCM delay and back of queue models in Eqs. (4) and (5).

$$
\begin{aligned}
d_1 &= \frac{C}{2}\left[\frac{(1-G/C)^2}{1-[min(1,X)G/C]}\right] \\
d_2 &= 900T\left[(X-1)+\sqrt{(X-1)^2+\frac{8kIX}{cT}}\right]
\end{aligned}
\tag{4}
$$

where d=$d_1 \times PF + d_2 + d_3$ is control delay seconds per vehicle, d_2 is uniform delay, PF is progression factor due to arrival types, d_2 is overflow delay component, and $d3$ is delay due to initial queue. In this study, only $d_1 + d_2$ are considered with $d_3 = 0$ since no clear queue growth was observed. $PF = 1.0$ was assumed for random arrivals. Volume-to-capacity $X = v/c = 0.5$ was assumed. Green time G is in seconds s, C is cycle time in s. T denoted analysis period in hours where in cycle-to-cycle estimations $T_i = C_i/3600$ was assumed where i denotes cycle number. k is incremental delay factor and 0.5 was assumed for pretimed like movement. $I = 1$ upstream filtering was assumed for no interaction with nearby intersections, and $c = 1029\,vph$ was assumed as explained above. Note that in our calculations, uniform delay is the main component updated by changing G and C values. Queue lengths were approximated by Little's formula $d \times \lambda$ where d and λ are both calculated at each cycle using M number of connected vehicles in the queues (Little, 1961). This method is based on HCM 2000 adopted from (Roess et al., 2004) and checked updated versions from (Prassas and Roess, 2020; Ni, 2020).

Moreover, a simpler approach from (Kyte and Tribelhorn, 2014) was used to calculate cycle-to-cycle back of queues which is given in Eq. (5).

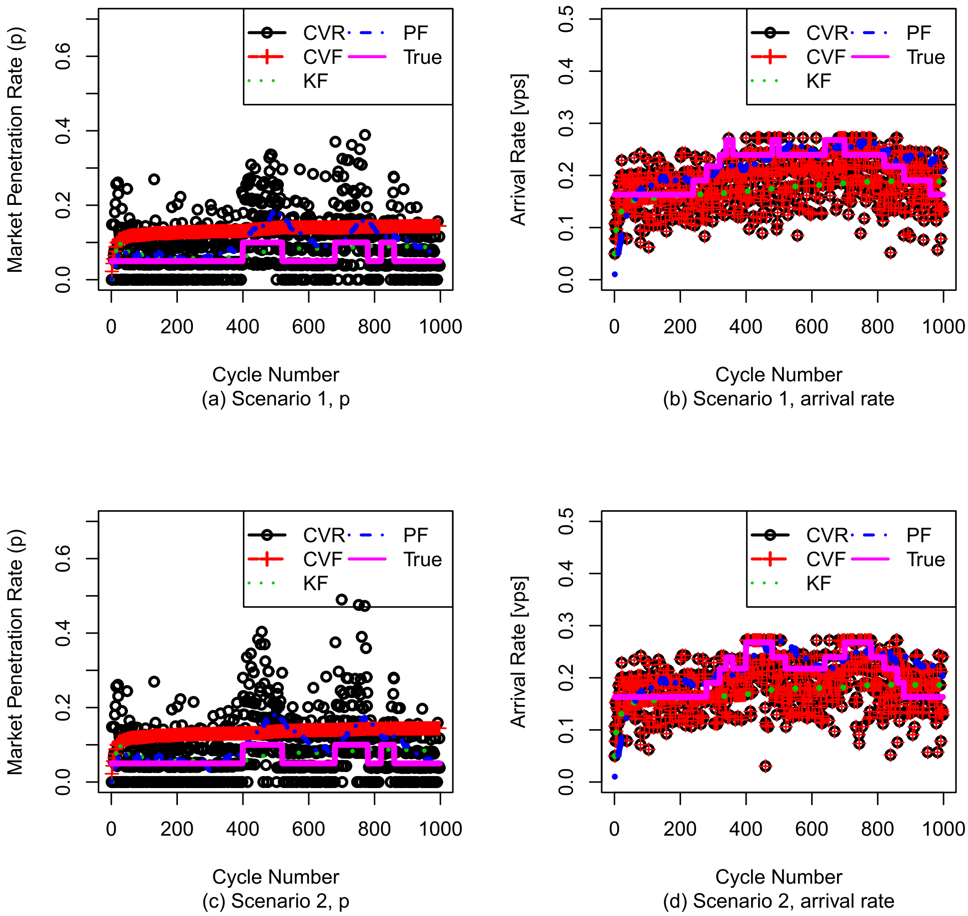

Fig. 10. Parameter estimation example with filters at $p\%$, $v/c = 0.99$ as average of 3 random seeds.

$$Qback = v(R + g_s) \tag{5}$$

where, $Qback$ is back of the queue in vehicles, v is arrival rate in vehicles per second (vps), R is red duration in seconds s, and g_s is queue service time which can be calculated $vR/(s - v)$ with s is saturation flow rate (i.e., assumed as 0.286 vps). All of the values (R, g_s, and v) except s are updated cycle-to-cycle.

Filtering approach was only showed using Kalman filter on using only M as estimator, $Avg = \frac{XM_i + \overline{M}}{2}$, $Est1 = \frac{M_i/R_i + C_i\overline{M}_{i-2:i}}{2}$, and $Est2 = \frac{C_i\left(XM_i/\overline{C}_{i-2:i}\right) + M_i/(XM_i/max(M_{1:i}))}{2}$ where notation $M_{1:i}$ represents values from cycle 1 to i and $\overline{M}_{1:i} = \sum_{j=1}^{i} \frac{M_j}{i}$.

In Table 1, summary of average queue length estimation errors are provided in root mean squared errors (RMSE=$\sqrt{\sum_{i=1}^{n} \frac{\left(QL_i - \hat{QL}_i\right)^2}{n}}$). Overall average market penetration rates for each lane are also given. Since, true maximum queues

Table 1
Back of queue estimation results with RMSE errors in [vehs/cycle].

	Lane	Avg. p	M	Avg	Est1	Est2	HCM	Q back	Imp.
Sep 08	L	13%	1.171	1.355	1.233	1.023	1.328	1.176	0.153
	C	21%	1.016	1.266	1.153	0.830	1.323	0.998	0.168
	R	7%	0.593	1.315	0.599	0.737	0.637	0.586	-0.151
Sep 09	L	10%	1.314	1.275	1.349	0.948	1.413	1.316	0.369
	C	26%	1.535	1.392	1.674	1.119	1.880	1.534	0.415
	R	2%	0.349	1.834	0.360	0.916	0.368	0.350	-0.566
Sep 10	L	7%	2.831	1.899	2.877	2.024	2.944	2.820	0.795
	C	18%	1.778	2.316	2.435	1.492	2.624	2.315	0.823
	R	1%	0.862	2.241	0.865	1.143	0.869	0.860	-0.283

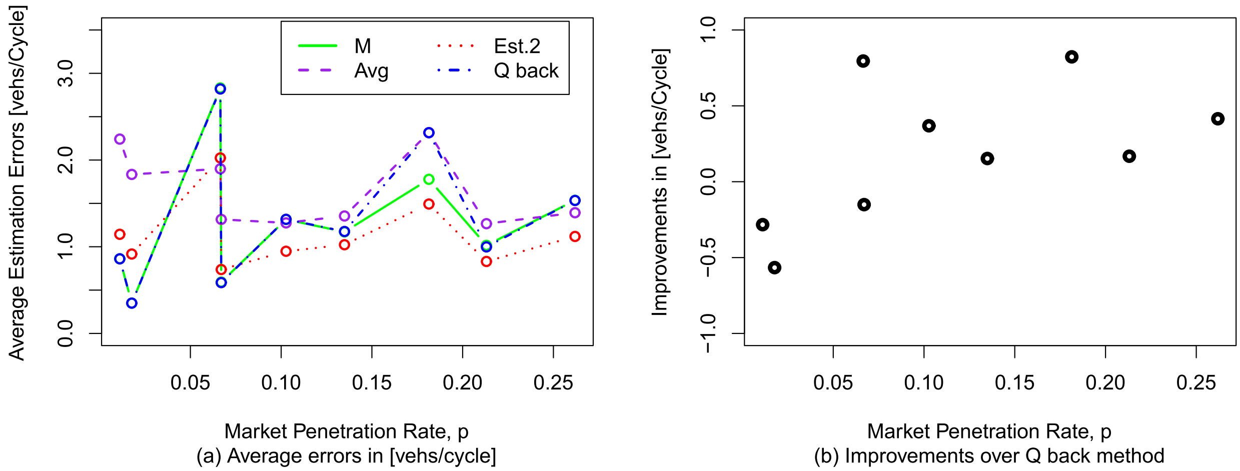

Fig. 11. Summary of queue length estimation results.

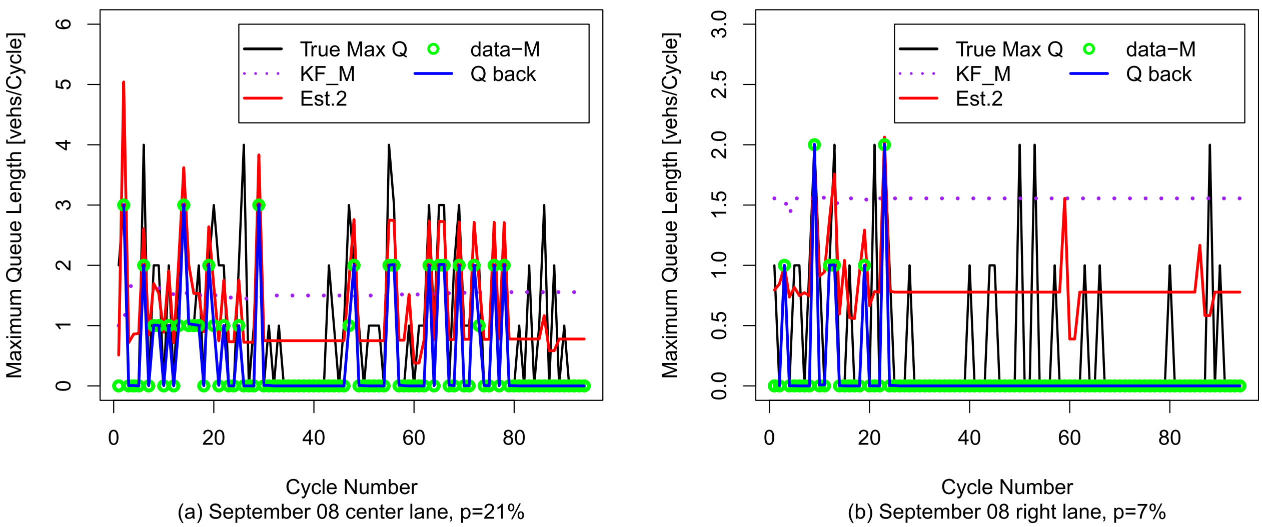

Fig. 12. Example queue length estimations on connected vehicle test data.

are not known, these *p*s are not used in estimations. Similar to simulation results, estimation with filtering is more accurate when some data is received from connected vehicles p=10%, 13%, 18%, 21%, *and*26%. Mixed results are obtained up to $p = 10\%$. In Fig. 11-a, -a, Est2 clearly performs better as p increase. Using only M for queue estimation is accurate when there is no vehicle in the queue at all cases. Having other information (L, T), these cases can be identified in Est2 and it could be improved. In Fig. 11-b, -b, trend of higher improvement of *Est*2 over *Qback* method with respect to p can be seen.

Example performance with high and low penetration rates are given in Fig. 12. When there is a probe vehicle in the queue, Est2 is able to produce more accurate estimation (Fig. 12-a). Kalman filter needs observations to be able to update estimates as it can be seen in the failures in right lane estimates at p=7% for series of zero queue values (Fig. 12-b).

5. Conclusions

In this study, we investigated the impact of localization framework on queue length estimation at low market penetration rates. We used two different filtering methods: (1) Kalman filter: unimodal and Gaussian noise and (2) Particle filter: multimodal and noise from any distribution.

The filters are fitted to queue length estimation framework where no true feedback is available. Two cases of fixed arrival and market penetration rate parameters and changing parameter are analyzed. Queue length estimation with filters show promising results improving accuracy by 20% for high volume-to-capacity ratio of $v/c = 0.99$. For lower v/c levels, with filtering by 30% market penetration rates, queue length estimation errors with known parameters (λ, p) are matched.

Estimation models and filters are both simple and enable real-time applications with no significant overhead. When there are some connected vehicles are present, filtering would be able to improve estimation. The approach can be used for real-time applications (e.g., signal control, ramp metering, and queue warning) or complement HCM methods to approximate control delay. In HCM methods, parameters are assumed to be unchanging in short-term intervals (e.g., 5 to 15-min) intervals where filtering can be used to locate true parameter values after low number of observations. Moreover, proposed approach can also be used if partial observations are available to better approximate the central tendency measures.

For future research, alternative queue length estimation models will be included for more realistic arrival and service time distributions as well as other intelligent transportation system applications.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study is partially supported by the Center for Connected Multimodal Mobility (C^2M^2) (USDOT Tier 1 University Transportation Center) headquartered at Clemson University, Clemson, South Carolina. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Center for Connected Multimodal Mobility (C^2M^2) and the official policy or position of the USDOT/OST-R, or any State or other entity, and the U.S. Government assumes no liability for the contents or use thereof. It is also partially supported by U.S. Department of Homeland Security Summer Research Team Program Follow-On grant, NSF Grant Nos. 1719501, 1436222, 1954532, 1400991, and NASA ULI project headquartered at University of South Carolina, Columbia.

References

Aljamal, M.A., Farag, M., Rakha, H.A., 2020. Developing data-driven approaches for traffic density estimation using connected vehicle data. IEEE Access 8, 219622–219631.
Argote, J., Christofa, E., Xuan, Y., Skabardonis, A., 2011. Estimation of measures of effectiveness based on connected vehicle data. In: 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 1767–1772.
Asanjarani, A., Nazarathy, Y., Pollett, P.K., 2017. Parameter and state estimation in queues and related stochastic models: A bibliography. arXiv preprint arXiv:1701.08338.
Bernardo, J., Bayarri, M., Berger, J., Dawid, A., Heckerman, D., Smith, A., West, M., 2011. Particle learning for sequential bayesian computation. Bayesian Stat. 9 (9), 317.
Carvalho, C.M., Johannes, M.S., Lopes, H.F., Polson, N.G., et al, 2010. Particle learning and smoothing. Stat. Sci. 25 (1), 88–106.
Comert, G., 2013. Effect of stop line detection in queue length estimation at traffic signals from probe vehicles data. Eur. J. Oper. Res. 226 (1), 67–76.
Comert, G., 2016. Queue length estimation from probe vehicles at isolated intersections: Estimators for primary parameters. Eur. J. Oper. Res. 252 (2), 502–521.
Comert, G., Cetin, M., 2020. Queue length estimation at traffic signals: Connected vehicles with range measurement sensors. arXiv preprint arXiv:2011.09397.
Dataset, C., 2014. ITS World Congress Connected Vehicle Test Bed Demonstration Vehicle Situation Data. South East Michigan Test Bed Contractor Team, Noblis' Queue Length Algorithm Development Team, and Data Capture and Management Data Sets Contractor Team, provided by ITS DataHub through Data.transportation.gov. Accessed 2021–01-25 from http://doi.org/10.21949/1504496.
Doucet, A., Johansen, A.M., 2009. A tutorial on particle filtering and smoothing: Fifteen years later. Handbook Nonlinear Filtering 12 (656–704), 3.
Emami, A., Sarvi, M., Bagloee, S.A., 2019. Using kalman filter algorithm for short-term traffic flow prediction in a connected vehicle environment. J. Modern Transp. 27 (3), 222–232.
Grumert, E.F., Tapani, A., 2018. Traffic state estimation using connected vehicles and stationary detectors. Journal of advanced transportation 2018.
Guo, Q., Li, L., Ban, X.J., 2019. Urban traffic signal control with connected and automated vehicles: A survey. Transp. Res. Part C: Emerging. technologies.
Gustafsson, F., Gunnarsson, F., Bergman, N., Forssell, U., Jansson, J., Karlsson, R., Nordlund, P.-J., 2002. Particle filters for positioning, navigation, and tracking. IEEE Trans. Signal Process. 50 (2), 425–437.
Khan, S.M., Dey, K.C., Chowdhury, M., 2017. Real-time traffic state estimation with connected vehicles. IEEE Trans. Intell. Transp. Syst. 18 (7), 1687–1699.
Kyte, M., Tribelhorn, M., et al, 2014. Operation, analysis, and design of signalized intersections: a module for the introductory course in transportation engineering Tech. rep.. TranLIVE. University of Idaho.
Labbe, R., 2014. Kalman and bayesian filters in python. Chap 7, 246.
Li, J.-Q., Zhou, K., Shladover, S.E., Skabardonis, A., 2013. Estimating queue length under connected vehicle technology: Using probe vehicle, loop detector, and fused data. Transp. Res. Record 2356 (1), 17–22.
Little, J.D., 1961. A proof for the queuing formula: $L = \lambda$ w. Operations Res. 9 (3), 383–387.
Liu, H., Liang, W., Rai, L., Teng, K., Wang, S., 2019. A real-time queue length estimation method based on probe vehicles in cv environment. IEEE Access 7, 20825–20839.
Liu, J.S., Chen, R., 1998. Sequential monte carlo methods for dynamic systems. J. Am. Stat. Assoc. 93 (443), 1032–1044.
Lopes, H.F., 2009. Tutorial: An introduction to particle filters. Unpublished Tutorial.
Lopes, H.F., Tsay, R.S., 2011. Particle filters and bayesian inference in financial econometrics. J. Forecasting 30 (1), 168–209.
Mandal, A., Sadhukhan, P., Gaji, F., Sharma, P., 2020. Measuring real-time road traffic queue length: A reliable approach using ultrasonic sensor. In: Proceedings of the 2nd International Conference on Communication, Devices and Computing. Springer, pp. 391–398.
Mei, Y., Gu, W., Chung, E.C., Li, F., Tang, K., 2019. A bayesian approach for estimating vehicle queue lengths at signalized intersections using probe vehicle data. Transp. Res. Part C: Emerging Technol. 109, 233–249.
Ni, D., 2020. Signalized Intersections. Springer.
Prassas, E.S., Roess, R.P., 2020. The Highway Capacity Manual: A Conceptual and Research History, Volume 2. Springer.
Roess, R.P., Prassas, E.S., McShane, W.R., 2004. Traffic engineering. Pearson/Prentice Hall.
Rompis, S.Y., Cetin, M., Habtemichael, F., 2018. Probe vehicle lane identification for queue length estimation at intersections. J. Intell. Transp. Syst. 22 (1), 10–25.

Tan, C., Liu, L., Wu, H., Cao, Y., Tang, K., 2020. Fuzing license plate recognition data and vehicle trajectory data for lane-based queue length estimation at signalized intersections. J. Intell. Transp. Syst., 1–18
Tan, C., Yao, J., Tang, K., Sun, J., 2019. Cycle-based queue length estimation for signalized intersections using sparse vehicle trajectory data. IEEE Trans. Intell. Transp. Syst.
Thrun, S., 2002. Particle filters in robotics. UAI 2, 511–518.
Thrun, S., Burgard, W., Fox, D., 2006. Probalistic robotics. Kybernetes.
Tiaprasert, K., Zhang, Y., Wang, X.B., Zeng, X., 2015. Queue length estimation using connected vehicle technology for adaptive signal control. IEEE Trans. Intell. Transp. Syst. 16, 2129–2140.
University of Arizona, University of California PATH Program, S.N.I.E., 2016. Multi-modal intelligent traffic signal system-phase ii: System development, deployment and field test.
Van Phu, C.N., Farhi, N., 2018. Estimation of urban traffic state with probe vehicles. arXiv preprint arXiv:1811.05394.
Wu, A., Yang, X., 2013. Real-time queue length estimation of signalized intersections based on rfid data. Procedia-Social Behav. Sci. 96, 1477–1484.
Xu, H., Ding, J., Zhang, Y., Hu, J., 2017. Queue length estimation at isolated intersections based on intelligent vehicle infrastructure cooperation systems. In: 2017 IEEE Intelligent Vehicles Symposium (IV). IEEE, pp. 655–660.
Yang, K., Menendez, M., 2018. Queue estimation in a connected vehicle environment: A convex approach. IEEE Trans. Intell. Transp. Syst. 20 (7), 2480–2496.
Yin, J., Sun, J., Tang, K., 2018. A kalman filter-based queue length estimation method with low-penetration mobile sensor data at signalized intersections. Transp. Res. Rec. 2672 (45), 253–264.
Zhao, Y., Zheng, J., Wong, W., Wang, X., Meng, Y., Liu, H.X., 2019. Estimation of queue lengths, probe vehicle penetration rates, and traffic volumes at signalized intersections using probe vehicle trajectories. Transp. Res. Rec. 2673 (11), 660–670.
Zhao, Y., Zheng, J., Wong, W., Wang, X., Meng, Y., Liu, H.X., 2019. Various methods for queue length and traffic volume estimation using probe vehicle trajectories. Transp. Res. Part C: Emerging Technol. 107, 70–91.