Applied Mathematics and Computation 421 (2022) 126915

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Robust estimation of traffic density with missing data using an adaptive-R extended Kalman filter

A.S.M. Bakibillah[a], Yong Hwa Tan[a], Junn Yong Loo[a], Chee Pin Tan[a,*], M.A.S. Kamal[b], Ziyuan Pu[a]

[a] School of Engineering and Advanced Engineering Platform, Monash University, Selangor, Bandar Sunway 47500, Malaysia
[b] Graduate School of Science and Technology, Gunma University, Kiryu 376–8515, Japan

ARTICLE INFO

Article history:
Received 7 August 2021
Revised 7 December 2021
Accepted 31 December 2021
Available online 14 January 2022

Keywords:
AREKF
Data imputation
Ramp metering
Traffic congestion
Traffic density estimation

ABSTRACT

Traffic density is a crucial indicator of traffic congestion, but measuring it directly is often infeasible and hence, it is usually estimated based on other measurements. However, a challenge in measuring traffic parameters is the high probability of sensor failure, which results in missing measurement or missing data. To overcome this difficulty, in this paper, we propose a novel adaptive-R extended Kalman filter (AREKF) combined with a model-based data imputation technique to estimate traffic density. We show analytically that the AREKF is able to accurately estimate the density even when the noise covariance matrices are not accurately known. Microscopic traffic simulations demonstrated the efficacy of the AREKF, where the estimated density is fed into a real-time ramp metering control algorithm to control vehicle flow on a merging road, which is highly susceptible to traffic congestion. The results show that the proposed AREKF with data imputation is able to accurately estimate the traffic density even when data is missing, and the ramp-metering controller significantly improves the traffic flow and thus, alleviates congestion.

1. Introduction

Traffic congestion is a major cause of energy consumption, greenhouse gas (GHG) emissions, and human stress. According to the INRIX Global Traffic Scorecard, drivers in the United States wasted 99 hours a year (on average) due to traffic congestion, costing about USD 88 billion in 2019, while drivers in the United Kingdom lost 115 hours (on average), costing about GBP 6.9 billion [1]. Traffic congestion can be divided into recurring and non-recurring [2]. The non-recurring types are usually temporary, due to environmental conditions, irregular individual driving styles or accidents. In contrast, recurrent congestion occurs regularly due to limited road capacity, merging and diverging traffic streams, infrastructure, and timing of traffic signals; their effects can (and should) be mitigated using intelligent transport systems (ITS), which yield congestion mitigation solutions using real-time traffic status, such as traffic volume, vehicle speed, and traffic density.

Traffic density, defined as the number of vehicles per unit length, is a critical parameter as it immediately shows the level of congestion and is frequently used to evaluate system-wide traffic performance; however, measuring it directly is not straightforward and often infeasible [3]. Traffic density is a range concept that can be characterized as continuous in space but discrete in time. One way of measuring density is using dense point sensors to approximate continuous measurements

* Corresponding author.
E-mail address: tan.chee.pin@monash.edu (C.P. Tan).

https://doi.org/10.1016/j.amc.2022.126915
0096-3003/

in space, but this is often too expensive. Another option is to take a snapshot of traffic along a stretch of a roadway [4]; however, this requires aerial cameras and real-time analysis, which is hardly available. It is also not suitable to obtain traffic density from the traffic occupancy rate of the road section using loop detectors, because the relationship is not constant and loop detectors are prone to substantial errors when placed close to the merging points [5]. Therefore, with these drawbacks in measuring traffic density, it is more practical to indirectly *estimate* traffic density, based on other variables that are easier to measure, such as speed and traffic flow.

Several approaches have been proposed to estimate traffic density. In [6], a cell transmission model (CTM) is used to estimate traffic density at a highway. However, this may be computationally intensive when combinations of small cells and various facility types are used to increase its accuracy. In [7], a fully-distributed vehicle-to-vehicle (V2V) communication-based method is proposed for real-time density estimation; however, V2V technology is difficult to practically implement. Coifman [8] used a sparse vehicle reidentification technique to estimate density and lane inflow; however, it is highly dependent on the accuracy of a loop detector-based vehicle re-identification algorithm and it did not consider a dynamic traffic model. In [9], a non-linear adaptive observer-based traffic density estimation scheme is proposed for real-time on-ramp metering control, but the estimation accuracy is poor when the traffic flow is congested.

One attractive option for estimating traffic density is using the Kalman filter, that can produce an optimal estimate based on measurements and a model that are noisy and uncertain; in the context of traffic systems, these uncertainties arise from traffic flow fluctuations and disruptions as well as measurement inaccuracies due to the harsh outdoor environment [10,11]. Singh and Li [12] developed an extended Kalman filter (EKF) to estimate traffic density using the information provided by loop detectors. Andre and Ny [13] proposed an ensemble Kalman filter (EnKF) for traffic density estimation along a road segment, while providing individual drivers with formal differential privacy guarantees. Di et al. [14] developed a hybrid EKF to estimate traffic density along a signalized arterial that uses data from both detectors and a global positioning system (GPS). In [15], a Lagrangian-space Kalman filter is proposed for freeway traffic density estimation based on the travel time transition model (TTM). In [16], an unscented Kalman filter (UKF) is developed for freeway traffic density estimation.

A drawback of Kalman filters is that they require knowledge on the covariance matrices of the process and measurement noises; this knowledge could be difficult to obtain. A way to overcome this is through Kalman filters that *adaptively* estimate the noise covariances; these have found applications in a variety of engineering domains. For example, Mohamed et al. [17] applied the adaptive Kalman filter to the trajectory tracking problem in a combined inertial navigation and global positioning system. Lunni et al. [18] used an adaptive extended Kalman filter to estimate the angle of rotation at each flexural joint of a tendon-driven soft robot. Sun et al. [19] used an adaptive unscented Kalman filter to estimate the state of charge of a lithium-ion battery for battery electric and hybrid electric vehicles. To the best of our knowledge, adaptive Kalman filters have not been applied to estimate traffic density, which is a crucial traffic parameter for depicting traffic status and is often used as the control system's input in a variety of scenarios, such as ramp metering control, signalized junction control, and so on. Due to the extremely dynamic nature of traffic flows in real-world scenarios, traffic systems are prone to numerous uncertainties and disturbances. Furthermore, there are measurement errors due to the high likelihood of sensor failure in harsh environments. As a result, it is necessary to use an adaptive Kalman filter to estimate real-world traffic density.

In this paper, to address the above-mentioned research gaps, we develop a novel traffic density estimation method based on an adaptive-R EKF (AREKF) using speed and traffic flow information which are easily measurable using standard traffic sensors. We propose a residual-based approach to adaptively estimate the covariance of the measurement noise (typically known as R). Moreover, we prove that the estimate of R is optimal with respect to the residual sequence when the posterior estimation error reaches steady state. As a result, it is no longer necessary to accurately know R, which is the covariance of the measurement noise. We assume that the speed data is always available, whereas traffic flow data contain missing values. This is a common occurrence with outdoor sensors exposed to harsh environmental conditions. To overcome the problem of missing traffic flow data, we incorporate a model-based data imputation technique. Based on the estimated traffic density (from the AREKF), we implement a real-time ramp metering strategy called ALINEA (which is the acronym for Asservissement line aire dentree autoroutie) to adaptively control the traffic flow of a real merging road. It is found that our proposed method can accurately estimate the traffic density (without accurately knowing R) compared to the standard extended Kalman filter (EKF) and can be implemented in real-time. Moreover, the ALINEA-based ramp metering algorithm that uses the estimated density significantly improves traffic flow performances of the merging road section under study.

This paper is organized as follows. In Section 2, we first describe the real study road stretch and data collection process, and the nonlinear traffic density model. Then we present the missing data imputation technique, the AREKF, and the real-time ramp metering control strategy. Section 3 illustrates key simulation results, and finally Section 4 gives the conclusion.

2. Methodology

2.1. Study road stretch and data collection

We consider a real on-ramp merging road section in Sunway City Kuala Lumpur, Malaysia as shown in Fig. 1(a). The main road, Jalan Lagoon Selatan (where 'Jalan' means 'road' in Malaysia) is a double-lane road and the merging (side) road Jalan PJS 11/9 is a single-lane road with moderate to congested traffic flow. The study road segment is 0.225 km long. To mimic

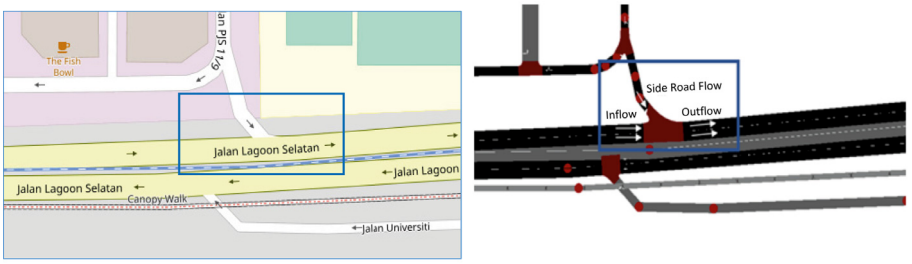

Fig. 1. (a) The study merging road section in Sunway City Kuala Lumpur, Malaysia taken from the Open Street Map. (b) Traffic flow network diagram in SUMO. The rectangles (blue) indicate the region of interest. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

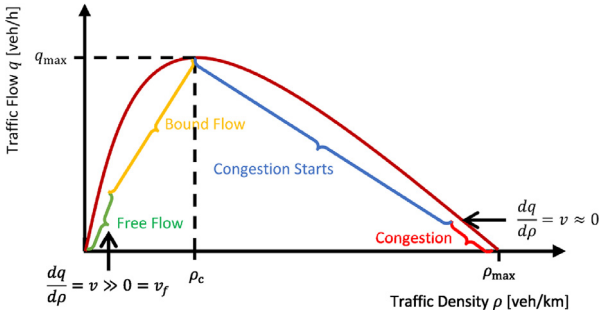

Fig. 2. Macroscopic fundamental diagram.

traffic flow behavior of the study road section, we use a simulator called the Simulation of Urban Mobility (SUMO), which is open source and used widely for discrete-time traffic simulations [20].

SUMO runs on the XML programming language and thus, the study road map is extracted in XML format (as an osm.xml file) from the Open Street Map. The map is converted into a network diagram (net.xml format) using the NETCONVERT function in SUMO as shown in Fig. 1(b). Then, the activity of traffic mobility (trips.xml) is generated using the ACTIVITYGEN function based on the population statistics of the study area. Then, the DUAROUTER function is used to generate the route for each vehicle based on the activity, which is stored in the rou.xml file. From SUMO, various macroscopic traffic parameters can be obtained, such as traffic density of the section, space mean speed, traffic inflow and outflow, and traffic flow from the side road to the main road. The traffic parameters obtained from SUMO are in XML format, which is not suitable for analysis. Hence, the data is converted into CSV format and MATLAB is used to process them. Using a TCP-based server/client architecture to access SUMO, the function Traffic Control Interface (TRACI) is used to facilitate communication between SUMO and MATLAB, where the data imputation method and AREKF are implemented. Then, the estimated density (from the AREKF) is used to control the ramp metering in real-time through TRACI.

We assume that the sensor reading of space mean speed is always available since the speed data is calculated as an average over a certain time interval. On the other hand, the traffic flow data from the sensor is recorded as a progressive number. Hence, the speed data is less affected by sensors malfunction compared to the flow data [21].

2.2. Nonlinear traffic model

In this paper, we develop a lumped parameter model to describe traffic flows and consider only temporal variation in the system variables. To develop the mathematical model, it is necessary to analyze traffic conditions using the macroscopic fundamental diagram (MFD), which is a visual representation of the state of road traffic (Fig. 2) and depicts the relation between traffic flow, density, and speed. It is commonly described as a triangle representing three states of traffic conditions, i.e., (i) the free flow of vehicles where the traffic density and flow rates are at their lowest, (ii) the critical density where traffic flow is at its highest (the maximum point of the curve) without slowing the traffic, and (iii) the maximum (jam) density where traffic congestion occurs; traffic flow is low and in severe circumstances, it is almost zero. From the MFD, the speed and traffic flow can be calculated as a function of density as

$$v_m(\rho(k)) = v_f \exp\left(-\left(\frac{1}{\alpha}\right)\left(\frac{\rho(k)}{\rho_c}\right)^{\alpha}\right), \tag{1}$$

$$q_m(\rho(k)) = \rho(k) v_m(\rho(k)), \tag{2}$$

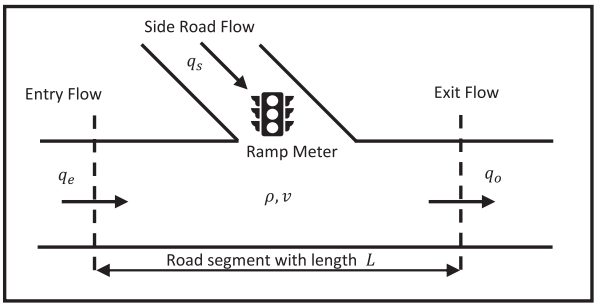

Fig. 3. Schematic representation of the study road segment. The traffic flow within the dashed line is of interest.

where $v_m(\rho(k))$, $q_m(\rho(k))$, and $\rho(k)$ respectively are the speed, traffic flow, and density at any time k, v_f is the free flow velocity, ρ_c is the critical density, and $\alpha \in \mathbb{R}_+$ is a positive constant. From the MFD, the corresponding speed $v_m(\rho(k))$ and flow $q_m(\rho(k))$ can be calculated if the density $\rho(k)$ is correctly known. At steady state, $q_m(0) = 0, q_m(\rho_{\text{jam}}) = 0, dq_m(\rho_c)/d\rho = 0$, and $v_m(0) = v_f, v_m(\rho_{\text{jam}}) = 0, dv_m(\rho_c)/d\rho \leq 0$, where ρ_{jam} is the jam density of the MFD. The MFD can be used to calculate the free flow velocity and critical density.

The on-ramp road section is schematically illustrated in Fig. 3. Let $N(k)$ be the number of vehicles in the road segment at time k. The conservation equation for vehicles in the segment with time step h is $N(k+1) = N(k) + h(q_e(k) - q_o(k) + q_s(k))$, where $q_e(k)$, $q_o(k)$, and $q_s(k)$ are the entry flow rate, exit flow rate, and side road (merge) flow rate of vehicles, respectively. The traffic density in the segment is the number of vehicles divided by its length L as

$$\rho(k+1) = \rho(k) + \frac{h}{L}(q_e(k) - q_o(k) + q_s(k)). \tag{3}$$

Next, the fundamental relation between flow q, density ρ, and speed v is $q = \rho v$. The gradient in the MFD (Fig. 2) is the speed v_m as a function of density ρ, which follows the fundamental relation. When the density $\rho(k)$ is accurately known, the actual speed $v(k)$ should be equal to the speed $v_m(\rho)$ from the MFD. Otherwise, if $\rho(k)$ is not correctly known, the error is given as

$$e(k) = v_m(\rho) - v(k). \tag{4}$$

We desire $e(k)$ to satisfy the following differential equation

$$\frac{de(k)}{dk} = -\alpha e(k), \tag{5}$$

whose solution is $e(k) = e(0)\exp(-\alpha k)$, where $e(0)$ is the initial error, which results in $e \to 0$. Then, substituting (4) into (5) and re-arranging gives

$$\frac{d(v_m(\rho) - v(k))}{dk} = \frac{dv_m(\rho)}{d\rho}\left(\frac{\partial \rho}{\partial k}\right) - \frac{dv(k)}{dk} = -\alpha(v_m(\rho) - v(k)). \tag{6}$$

After discretizing with time step h, (6) can be written as

$$\frac{v(k+1) - v(k)}{h} = \alpha(v_m(\rho) - v(k)) + \frac{dv_m(\rho)}{d\rho}\frac{\rho(k+1) - \rho(k)}{h}, \tag{7}$$

and substituting (3) into (7) gives

$$v(k+1) = v(k) + \alpha h(v_m(\rho) - v(k)) + \frac{dv_m(\rho)}{d\rho}\left(\frac{h}{L}\right)(q_e(k) - q_o(k) + q_s(k)). \tag{8}$$

By substituting (1) into (8), the space mean speed can be expressed as

$$\begin{aligned} v(k+1) = v(k) &+ \alpha h\left(v_f \exp\left(-\left(\frac{1}{\alpha}\right)\left(\frac{\rho(k)}{\rho_c}\right)^{\alpha}\right) - v(k)\right) \\ &- \frac{h v_f \rho(k)^{\alpha-1}}{L\rho_c{}^{\alpha}} \exp\left(-\left(\frac{1}{\alpha}\right)\left(\frac{\rho(k)}{\rho_c}\right)^{\alpha}\right)(q_e(k) - q_o(k) + q_s(k)). \end{aligned} \tag{9}$$

Next, we describe the missing data imputation technique and a discrete-time state-space representation with nonlinear process model for our traffic system.

2.3. Missing data imputation

To overcome the issue of missing data of traffic flow described above, we use a model-based missing data imputation technique, that replaces missing data with a reasonable value to reduce the inaccuracies due to the missing data. For each input ($q_e(k)$, $q_s(k)$, and $q_o(k)$), a corresponding NLARX (nonlinear autoregressive with exogenous input) function ($\hat{q}_e(v(k))$, $\hat{q}_s(v(k))$, and $\hat{q}_o(v(k))$, respectively) is generated and used to predict the missing data based on the measurement of space mean speed $v(k)$. Then, the outliers (caused by external influences) are removed from the data, and the input/output delay, the initial condition, and the number of poles and zeros are specified. After that, the data are divided into two parts; the first half of the data is used to estimate the model parameters and the second half to validate the model and analyze the residual. The process is repeated until a satisfactory model is generated.

2.4. State-space representation

We consider the following discrete-time state-space system

$$x(k+1) = f(x(k)) + w(k), \tag{10}$$

$$y(k) = Hx(k) + v(k), \tag{11}$$

where $x(k)$ is the system state, $w(k)$ is the process noise, $y(k)$ is the measurable output, and $v(k)$ is the measurement noise. $w(k)$ and $v(k)$ are assumed to be white noises i.e., they are zero-mean, uncorrelated ($\mathbb{E}[w(i)w(j)^T] = 0$, $\mathbb{E}[v(i)v(j)^T] = 0$ for $i \neq j$) and mutually independent ($\mathbb{E}[w(k)v(k)^T] = 0$). The nonlinear function f is a process model that propagates the state, whereas H is the measurement matrix that is constant.

By choosing traffic density (3) and space mean speed (9) as the state variables $x(k) = \begin{bmatrix}\rho(k) & v(k)\end{bmatrix}^T$, and since the space mean speed is measured, therefore $y(k) = v(k)$ and (10)-(11) can be re-written as

$$\underbrace{\begin{bmatrix}\rho(k+1)\\ v(k+1)\end{bmatrix}}_{x(k+1)} = \underbrace{\begin{bmatrix} \rho(k) + \frac{h}{L}\left(\hat{q}_e(v(k)) - \hat{q}_o(v(k)) + \hat{q}_s(v(k))\right) \\ v(k) + \alpha h\left(v_f \exp\left(-\left(\frac{1}{\alpha}\right)\left(\frac{\rho(k)}{\rho_c}\right)^{\alpha}\right) - v(k)\right) \\ -\frac{h v_f \rho(k)^{\alpha-1}}{L\rho_c{}^{\alpha}} \exp\left(-\left(\frac{1}{\alpha}\right)\left(\frac{\rho(k)}{\rho_c}\right)^{\alpha}\right) \\ \times\left(\hat{q}_e(v(k)) - \hat{q}_o(v(k)) + \hat{q}_s(v(k))\right) \end{bmatrix}}_{f(x(k))} + w(k), \tag{12}$$

$$y(k) = \underbrace{\begin{bmatrix}0 & 1\end{bmatrix}}_{H}\underbrace{\begin{bmatrix}\rho(k)\\ v(k)\end{bmatrix}}_{x(k)} + v(k), \tag{13}$$

with the aforementioned (known) parameters h, L, v_f, ρ_c , and α. Using the state-space model (12)-(13) as well as the measurement of $v(k)$, we can estimate the traffic density $\rho(k)$ via an adaptive-R extended Kalman filter detailed in the next section.

2.5. Adaptive-R extended Kalman filter

The Kalman filter is a state estimator that can recursively estimate the state of a dynamic system using its measurement. In this paper, we seek to estimate the density $\rho(k)$ based on the available data on the measurement i.e., the space mean speed $y(k) = v(k)$ in (13). The Kalman filter assumes that the process disturbance and measurement noise are normally distributed with zero mean and independent of each other. However, our process model (12) is nonlinear, and therefore, in this paper, we propose an adaptive-R extended Kalman filter (AREKF). The adaptive-R feature introduced in this paper enables adaptive estimation of the matrix R (which is the covariance matrix of the measurement noise $v(k)$); this is highly advantageous as the measurement noise covariance is not known and may vary over time (especially in traffic applications). With this adaptive-R feature, it releases the potentially restrictive assumption that the measurement noise covariance R is constant and known.

At each time step, the AREKF linearizes the nonlinear governing model and updates R. Taking the Taylor series expansion of (10) about $x = \hat{x}^+(k)$ yields

$$\begin{aligned} x(k+1) &= f(\hat{x}^+(k)) + F(k)(x(k) - \hat{x}^+(k)) + w(k), \\ &= F(k)x(k) + \tilde{u}(k) + w(k), \end{aligned} \tag{14}$$

where $\tilde{u}(k) = f(\hat{x}^+(k)) - F(k)\hat{x}^+(k)$ and $\hat{x}^+(k)$ is the posterior state estimation at time step k. $F(k)$ is Jacobian of the nonlinear model f w.r.t. x i.e., $F(k) = \frac{\partial f}{\partial x}|_{\hat{x}^+(k)}$.

With the Taylor series expansion above, we can then apply the adaptive Kalman filter equations which consist of a *prediction phase* and a *correction phase* on the augmented linear process model (14) as follows:

Prediction phase

1. The prior state estimation $\hat{x}^-(k+1)$ at time $k+1$ is predicted using the nonlinear process model $\hat{x}^-(k+1) = f(\hat{x}^+(k))$.
2. The prior state covariance $P^-(k+1)$ at time $k+1$ is predicted using the process disturbance covariance $Q(k)$ and the posterior state covariance $P^+(k)$ as follows $P^-(k+1) = F(k)P^+(k)F(k)^T + Q(k)$.

After the prediction phase, the AREKF adjusts the Kalman gain in the correction phase given below.

Correction phase

1. The difference between $y(k)$ and prior measurement estimation $\hat{y}^-(k) = H\hat{x}^-(k)$ gives the innovation $d(k)$ as $d(k) = y(k) - \hat{y}^-(k)$.
2. The covariance matrix $S(k)$ of the prior measurement $\hat{y}^-(k)$ and the Kalman gain $K(k)$ at time k are calculated as $S(k) = HP^-(k)H^T + R(k)$ and $K(k) = P^-(k)H^TS(k)^{-1}$.
3. The posterior state $\hat{x}^+(k)$ and posterior covariance $P^+(k)$ are calculated as $\hat{x}^+(k) = \hat{x}^-(k) + K(k)d(k)$ and $P^+(k) = P^-(k) - K(k)S(k)K(k)^T$.
4. The difference between $y(k)$ and posterior measurement estimation $\hat{y}^+(k) = H\hat{x}^+(k)$ gives the residual $r(k)$ as $r(k) = y(k) - \hat{y}^+(k)$.
5. The measurement noise covariance $R(k)$ at time step k can be adaptively approximated by the following residual-based time-weighted estimator

$$R(k) = \sum_{i=k-l+1}^{k} \omega(i)\left(r(i)r(i)^T + HP^+(i)H^T\right), \tag{15}$$

with the positive weights $0 \le \omega(i) \le 1$. In this paper, we let $\omega(i) = (1-\beta)\beta^{k-i}$ for $i = 1, \ldots, k$ and $\omega(0) = \beta^k$ with a forgetting factor $0 \le \beta \le 1$ and replace the moving window average (15) with

$$\begin{aligned} R(k) = \quad & \sum_{i=1}^{k}(1-\beta)\beta^{k-i}\left(r(i)r(i)^T + HP^+(i)H^T\right) \\ & +\beta^k\left(r(0)r(0)^T + HP^+(0)H^T\right). \end{aligned} \tag{16}$$

Additionally, (16) can be written recursively as

$$R(k) = \beta R(k-1) + (1-\beta)\left(r(k)r(k)^T + HP^+(k)H^T\right), \tag{17}$$

which is much more computationally efficient than (15) and (16) as it only requires the memory of the past estimate $R(k-1)$. The choice of β will be discussed in Remark 1 below.

Remark 1. Notice that the set of weights $(1-\beta, (1-\beta)\beta, (1-\beta)\beta^2, \ldots, (1-\beta)\beta^{k-1}, \beta^k)$ from (16) decrease exponentially and form a geometric progression of unit sum. Hence, β can be chosen such that the weights on $i \le k-l$ are negligible and (16) provides a good approximation of the moving window average (15). A larger β puts more weight on the past history and causes less fluctuation on the estimated $R(k)$, at the cost of slower convergence. In the case of a highly nonlinear system, however, β should be set small albeit cautiously because an excessively small β results in a heavier fluctuation to the estimated $R(k)$. Further reasons for choosing β to be small can be found in Remark 3. In this paper, we choose $\beta = 0.7$ from validating the estimation results on the first half of the data. Since the computation of posterior state $\hat{x}^+(k)$ requires $R(k)$ beforehand, we also made an one-step ahead approximation $R(k+1) \approx R(k)$, which is shown to be accurate at a small sampling time step of $h = 0.0833$ s for our traffic system.

Prior to the derivation of the residual-based R-estimator (15), we first present the following proposition that state the conditions for the residual $r(k)$ to be a white noise process.

Proposition 1. *Let $\hat{x}^+(k)$ be an unbiased estimate which satisfies the orthogonality condition $\mathbb{E}[(x(k) - \hat{x}^+(k))\psi(y(j))] = 0$ for $j \le k$ and any vector function $\psi(\cdot)$. Then the residual sequence $r(k)$ is a white noise process.*

Proof. Let $\tilde{x}(k) = x(k) - \hat{x}^+(k)$ denote the posterior state error. We expand the residual $r(k) = y(k) - \hat{y}^+(k)$ to get

$$\begin{aligned} r(k) \quad & = Hx(k) + \nu(k) - H\hat{x}^+(k) \\ & = H\tilde{x}(k) + \nu(k) \end{aligned} \tag{18}$$

For $i > j$, we have $\mathbb{E}[\nu(i)\nu(j)^T] = 0$ since $\nu(k)$ is uncorrelated, and $\mathbb{E}[\nu(i)\tilde{x}(j)^T] = 0$ because $\nu(i)$ is independent of $\tilde{x}(j)$ of previous time-steps. Hence, the residual covariance becomes

$$\begin{aligned} \mathbb{E}\left[r(i)r(j)^T\right] \quad & = \mathbb{E}\left[(H\tilde{x}(i) + \nu(i))(H\tilde{x}(j) + \nu(j))^T\right] \\ & = \mathbb{E}\left[H\tilde{x}(i)(H\tilde{x}(j) + \nu(j))^T\right] \\ & = \mathbb{E}\left[H\tilde{x}(i)(y(j) - H\hat{x}^+(j))^T\right] \end{aligned} \tag{19}$$

Applying the orthogonality condition on (19), it yields $\mathbb{E}[r(i)r(j)^T] = 0$ since $y(j) - H\hat{x}^+(j)$ is a function of $y(j)$. Repeating the steps for $i < j$ yields the same result $\mathbb{E}[r(i)r(j)^T] = 0$. Furthermore, we have $\mathbb{E}[r(k)] = 0$ due to the assumptions that $\hat{x}^+(k)$ is unbiased i.e., $\mathbb{E}[\tilde{x}(k)] = 0$ and $v(k)$ is zero-mean. Hence, $r(k)$ satisfies the definition of a white noise process. □

Remark 2. The unbiasness and orthogonality condition follows if the filter estimate $\hat{x}^+(k)$ is optimal i.e., $\hat{x}^+(k)$ is given by the expectation $\mathbb{E}[x(k)|Y(k)]$ conditioned on the measurement process $Y(k) = (y(0), y(1), \ldots, y(k))$ [22]. Due to the well-known moment closure problem, a finite-dimensional analytical solution to the conditional expectation exist only in the case of a linear system, which is given by the Kalman filter. The EKF is known to be a sub-optimal filter that approximates the conditional expectation in the case of a nonlinear system [23].

A residual-based *R*-estimator was first derived in [24] where it is derived to minimize the square of the innovation $d(k) = y(k) - \hat{y}^-(k)$. Proposition 2 below presents a more concise derivation by minimizing the time-weighted square of residual $r(k) = y(k) - \hat{y}^+(k)$, which leads the to the *R*-estimator in (15). Then, we show that (15) is an ergodic estimation to the actual value of $R(k)$ when the posterior state error $\tilde{x}(k) = x(k) - \hat{x}^+(k)$ is in steady state.

Proposition 2. *Assume that the posterior state error $\tilde{x}(k) = x(k) - \hat{x}^+(k)$ is in steady state and that the residual $r(k)$ is a Gaussian distributed white noise process. Then the Eq. (15) gives an maximum likelihood estimation or equivalently a minimum mean-square estimation to $R(k)$.*

Proof. Given that the residual $r(i)$ is a Gaussian white noise process on a time window $i = k - l + 1, \ldots, k$ of fixed length l, we have

$$p(r(k-l+1), r(k-l+2), \ldots, r(k)) = \prod_{i=k-l+1}^{k} p(r(i)), \tag{20}$$

with the Gaussian density function

$$p(r(i)) = \frac{1}{\sqrt{(2\pi)^m |C_{r(i)}|}} e^{-\frac{1}{2} r(i)^T C_{r(i)}^{-1} r(i)}, \tag{21}$$

where the log likelihood of $r(i)$ is of the form of $\ln p(r(i)) = -\frac{1}{2}(m \ln(2\pi) + \ln|C_{r(i)}| + r(i)^T C_{r(i)}^{-1} r(i))$.

To obtain an estimate of $R(k)$, we solve the following time-weighted maximum log likelihood problem w.r.t. the diagonal elements of $R(i)$, $R_{jj}(i)$ on the time window $i = k - l + 1, \ldots, k$,

$$\underset{R_{jj}(k-l+1)}{arg\,max} \ldots \underset{R_{jj}(k)}{arg\,max} \sum_{i=k-l+1}^{k} \omega(i) \ln p(r(i)). \tag{22}$$

with positive weights $0 \le \omega(i) \le 1$. This is equal to solving the following time-weighted generalized least square problem

$$\underset{R_{jj}(k-l+1)}{arg\,min} \ldots \underset{R_{jj}(k)}{arg\,min} \sum_{i=k-l+1}^{k} \omega(i) \left(\ln|C_{r(i)}| + r(i)^T C_{r(i)}^{-1} r(i) \right), \tag{23}$$

after switching signs and neglecting the constant terms. Each minimization solution w.r.t. $R_{jj}(i)$ can be obtained via taking the derivative and equating to zero, i.e.,

$$\sum_{i=k-l+1}^{k} \omega(i) \frac{\partial \left(\ln|C_{r(i)}| + r(i)^T C_{r(i)}^{-1} r(i) \right)}{\partial R_{jj}(i)} = 0, \tag{24}$$

which results in

$$\begin{aligned} \sum_{i=k-l+1}^{k} \omega(i) \mathrm{tr}\left[C_{r(i)}^{-1} \tfrac{\partial C_{r(i)}}{\partial R_{jj}(i)} \right] - r(i)^T C_{r(i)}^{-1} \tfrac{\partial C_{r(i)}}{\partial R_{jj}(i)} C_{r(i)}^{-1} r(i) = 0, \\ \sum_{i=k-l+1}^{k} \omega(i) \mathrm{tr}\left[\left(C_{r(i)}^{-1} - C_{r(i)}^{-1} r(i) r(i)^T C_{r(i)}^{-1} \right) \tfrac{\partial C_{r(i)}}{\partial R_{jj}(i)} \right] = 0, \end{aligned} \tag{25}$$

where we have used the linear algebra relation ($x^T A^T B A x = \mathrm{tr}[A x x^T A^T B]$) with the trace operator $\mathrm{tr}[\cdot]$, and the following matrix calculus identities

$$\frac{\partial \ln|A|}{\partial x} = \frac{1}{|A|} \frac{\partial |A|}{\partial x} = \mathrm{tr}\left[A^{-1} \frac{\partial A}{\partial x} \right], \quad \frac{\partial A^{-1}}{\partial x} = -A^{-1} \frac{\partial A}{\partial x} A^{-1} \tag{26}$$

where x is a vector and A, B are square matrices. The residual covariance matrix can also be written as [23]

$$C_{r(i)} = R(i) - HP^+(i)H^T. \tag{27}$$

The partial derivative of $C_{r(i)}$ w.r.t. $R(i)$ is then

$$\frac{\partial C_{r(i)}}{\partial R_{jj}(i)} = H \frac{\partial P^+(i)}{\partial R_{jj}(i)} H^T + \frac{\partial R(i)}{\partial R_{jj}(i)}. \tag{28}$$

Given the assumption that the estimation error $\tilde{x}(i) = x(i) - \hat{x}^+(i)$ is in steady state on the time window $i = k-l+1, \ldots, k$, then $P^+(i)$ is constant and $\frac{\partial P^+(i)}{\partial R_{jj}(i)} = 0$. Hence, (25) is reduced to

$$\sum_{i=k-l+1}^{k} \omega(i)\left[C_{r(i)}^{-1} - C_{r(i)}^{-1} r(i) r(i)^T C_{r(i)}^{-1}\right]_{jj} = 0, \quad j = 1, 2, \ldots, m \tag{29}$$

which is equivalent to

$$\sum_{i=k-l+1}^{k} \omega(i) \mathrm{tr}\left[C_{r(i)}^{-1}\left(C_{r(i)} - r(i) r(i)^T\right) C_{r(i)}^{-1}\right] = 0. \tag{30}$$

Note that the trace vanishes only if every entry of the matrix in the inner bracket equates to zero. Hence, we have

$$\sum_{i=k-l+1}^{k} \omega(i)\left(C_{r(i)} - r(i) r(i)^T\right) = 0. \tag{31}$$

Then, substitute (27) into (31) to get

$$\sum_{i=k-l+1}^{k} \omega(i) R(i) = \sum_{i=k-l+1}^{k} \omega(i)\left(r(i) r(i)^T + H P^+(i) H^T\right). \tag{32}$$

Given that the window length l is sufficiently long, invoking the ergodicity of the white measurement noise $\nu(k)$, we have

$$R(k) = \frac{1}{l} \sum_{i=k-l+1}^{k} \omega(i)\left(r(i) r(i)^T + H P^+(i) H^T\right) \tag{33}$$

which yields the R-estimator in (15). □

Remark 3. In the case of a highly nonlinear system (a large $\tilde{u}(k)$ for (14)), the state estimation error $\tilde{x}(k)$ converges slowly [25] and remains transient for a longer time; thus the time window length l in (15) should be set small to yield an approximately steady $\tilde{x}(k)$ so that $\frac{\partial P^+(i)}{\partial R_{jj}(i)} \approx 0$ within a short time window $i = k-l+1, \ldots, k$. Equivalently, we can set a small forgetting factor β in (16) or (17) so that the weights on $i \le k-l$ are negligible.

Remark 4. The Gaussian assumption on the residual sequence $r(k)$ is not mandatory and consequently the process and measurement noises $w(k)$ and $\nu(k)$ do not need to be Gaussian. This is because generalized least square is a standard objective function that yields an unbiased and minimum variance parameter estimation, regardless of the distribution of its residual [26].

Remark 5. In this paper, we consider only adaptively estimating R (known as the adaptive-R approach) as opposed to simultaneously estimating (adaptively) both Q and R. An obvious benefit of estimating only R is its reduced computational load, along with the other benefits that are expounded upon in the remainder of this remark. Though some recent works have simultaneously estimated both Q and R [27–30], their Q-estimator was derived by assuming prior knowledge of R, whereas their R-estimator was derived by assuming prior knowledge of Q. In fact, due to the recursive nature of the Kalman filter, the solution of the Q-estimator depends on the solution of the R-estimator and vice versa. Hence, current works that simultaneously estimate Q and R using separate estimators could result in non-unique and often-times biased estimates that do not converge to the actual values; the contribution of Q and R (at each time step) on the measurement covariance cannot be accurately distinguished from each other.

Moreover, the Q estimation in these works do not guarantee positive definite solutions that ensure numerical stability. A conventional Q-estimator (for example in [28]) that uses the average estimation of the innovation covariance $C_{d(k)} = \frac{1}{l}\sum_{i=k-l+1}^{k} d(i) d(i)^T$ involves matrix subtraction which could lead to a non-positive definite solution [31]. Conversely, the proposed residual-based R-estimator in (15) guarantees positive definite solutions.

Mehra [32] and Maybeck [24] have shown that (in the linear system analogue) unique and independent solutions of Q and R can be guaranteed only when there are equal or more measurements than states (measurement matrix H is of rank $\ge n$), which does not apply for many practical dynamic systems, including the traffic system considered in our paper, which is the reason that we adaptively estimate only R.

2.6. Ramp metering control

The imputed missing data and resulting density estimates will then be used to control the traffic flow of a merging road to prevent undesirable traffic phenomena, such as congestion and shockwaves [33]. In this paper, we use *ramp metering*, which is an effective strategy for reducing congestion on merging roadways [34], and we choose a ramp-metering technique known as ALINEA (the acronym for Asservissement line aire dentree autoroutie), which is simple, flexible, robust, and effective to adjust the ramp metering rate [35]. It ensures that the downstream traffic density (that is beyond the merging point on a highway) is close to the critical density by maximizing freeway throughput. The control rate of ALINEA is defined as

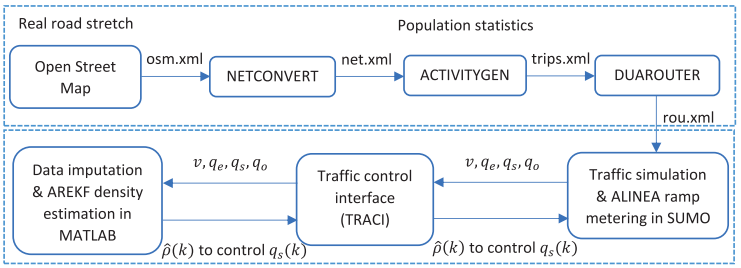

Fig. 4. Framework of traffic density estimation using the AREKF and ALINEA ramp metering control.

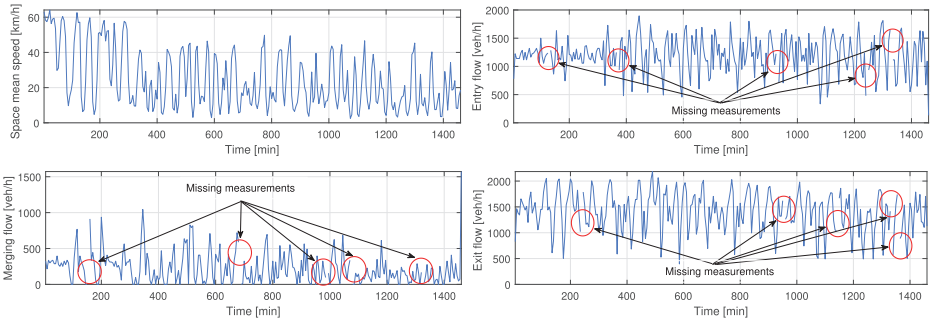

Fig. 5. Space mean speed $v(k)$ and missing measurements of entry flow $q_e(k)$, merge flow $q_s(k)$, and exit flow $q_o(k)$ due to sensor failure.

$q_s(k+1) = q_s(k) + \lambda_r\left(\rho_c - \hat{\rho}(k)\right)$, where $q_s(k)$ is the entrance ramp traffic flow at time k, λ_r is a controller parameter, and $\hat{\rho}(k)$ is the estimate of the density from the AREKF. Increasing λ_r leads to stronger regulator response and shorter regulatory time, and viceversa. For high values of λ_r, the regulator may have an oscillatory behavior or even worse, become unstable. If the estimated density $\hat{\rho}(k)$ is higher than the critical density ρ_c, the ramp meter rate $q_s(k)$ will decrease and viceversa; this strategy is applicable to both free-flow and congested traffic.

The overall concept of our proposed framework for traffic density estimate using the AREKF and ALINEA ramp metering control is depicted in Fig. 4.

3. Simulation results and discussion

We conducted microscopic traffic simulations considering the real on-ramp merging road section in Sunway City Kuala Lumpur, Malaysia (described in Section 2.1) using SUMO and collected the data on space mean speed $v(k)$, entry flow $q_e(k)$, merge (side road) flow $q_s(k)$, exit flow $q_o(k)$ and signal phases. The data are then processed using MATLAB to obtain information on space mean speed and traffic flows with missing measurements or missing data as shown in Fig. 5. Note that we considered several conditions of traffic flow and missing measures, e.g., entry, merge, and exit flows and durations for the density estimation. The missing measurements of traffic flow are generated randomly using the *randperm* function in MATLAB. The processed information is used to estimate traffic density with missing data using the model-based data imputation technique in Section 2.3 and AREKF in Section 2.4. Then, the estimated traffic density is applied to adaptively control the ramp metering in real-time.

3.1. Evaluation of MFD

The free flow velocity v_f, critical density ρ_c, and positive constant α parameters are tuned to fit the MFD to the data using the optimization toolbox. Fig. 6 illustrates the MFD of the studied road section. Since the fitted triangular curve matches the simulated traffic density better than the parabolic curve, it can be inferred that the triangular curve can better reflect real-world situations [36]. From the MFD, it is found that the road is highly congested, with the majority of traffic density data points exceeding the critical density ρ_c. The MFD yields the best-fit constant parameters for the nonlinear model, with v_f, ρ_c, and α being 92.84 veh/h, 51.32 veh/km, and 2, respectively.

3.2. Performance of AREKF

To estimate $\rho(k)$, the data on $v(k)$, $q_e(k)$, $q_s(k)$, and $q_o(k)$ are injected into the AREKF, and the estimate $\hat{\rho}(k)$ is obtained. The initial conditions of the AREKF are set as $x(0) = \begin{bmatrix}0 & 60\end{bmatrix}^T$, $P(0) = 10I_2$, and $R(0) = 100$. The time step h and process

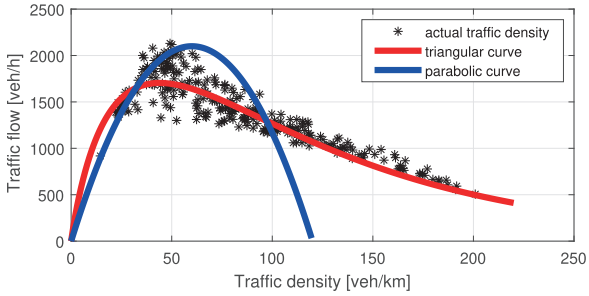

Fig. 6. MFD of the study merging road section.

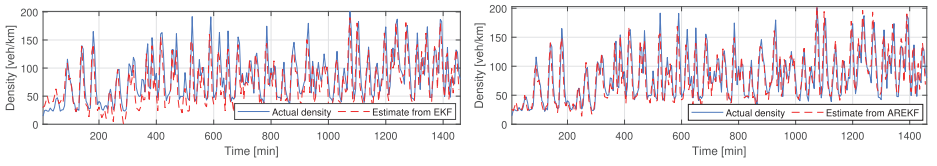

Fig. 7. Comparison of actual traffic density and estimated density from (a) the EKF and (b) the AREKF with normal data.

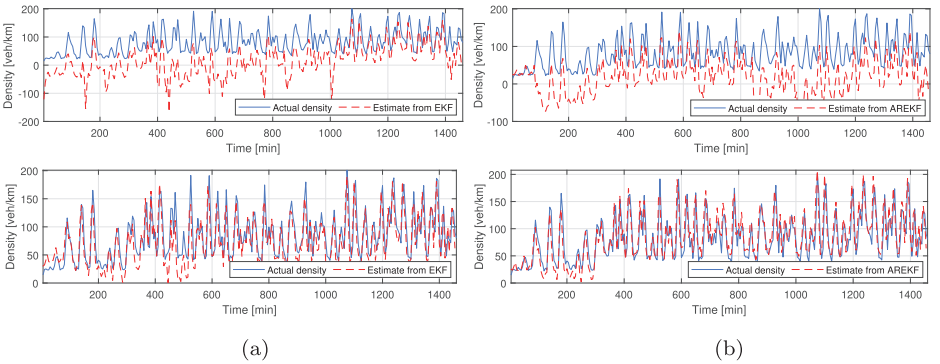

Fig. 8. Comparison of actual traffic density and estimated density from the (a) the EKF and (b) the AREKF with entry road missing data but without data imputation (top) and missing data imputation (bottom).

Table 1
Performance of AREKF and EKF for traffic density estimation .

Performance cases	AREKF MAPE Time [s]	EKF MAPE Time [s]
Normal data	21.53% 0.40	26.76% 0.38
Entry missing data	95.72% 0.42	109.49% 0.41
Impute entry missing data	26.04% 0.45	27.64% 0.44
Merging missing data	34.78% 0.41	40.85% 0.40
Impute merging missing data	23.62% 0.44	26.39% 0.43
Exit missing data	246.83% 0.42	373.98% 0.41
Impute exit missing data	27.45% 0.47	29.98% 0.45

disturbance covariance Q are set as 0.0833 s and $200I_2$, respectively. The proposed method is evaluated in terms of accuracy and computational time using *normal data* (when all sensors are functioning and no data is missing) as shown in Fig. 7 and several cases of traffic conditions and *missing data* (with and without missing data imputation) of entry, merge, and exit flows as shown in Figs. 8, 9, and 10, respectively, and the results are compared to the standard (non-adaptive) EKF.

The accuracy of the model is evaluated using the mean absolute percentage error (MAPE) as follows: $MAPE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{A_k - F_k}{A_k}\right|$, where A_k is the actual density, F_k is the estimated density, and n is the number of samples and justified using Lewis' scale of judgment for forecasting accuracy [37]. The overall performance of AREKF compared to the EKF in different cases is given in Table 1, which shows the accuracy of the AREKF with normal data and missing data imputation and considered accurate according to Lewis' scale of judgment of forecasting accuracy. The traffic density estimation with

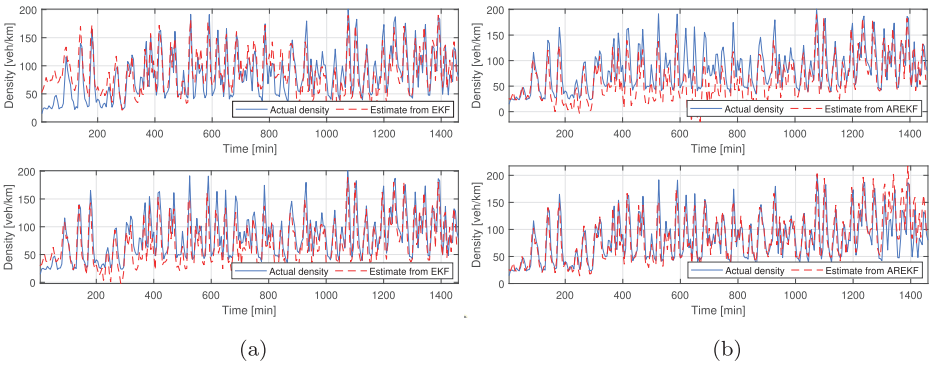

Fig. 9. Comparison of actual traffic density and estimated density from the (a) the EKF and (b) the AREKF with side road missing data but without data imputation (top) and missing data imputation (bottom).

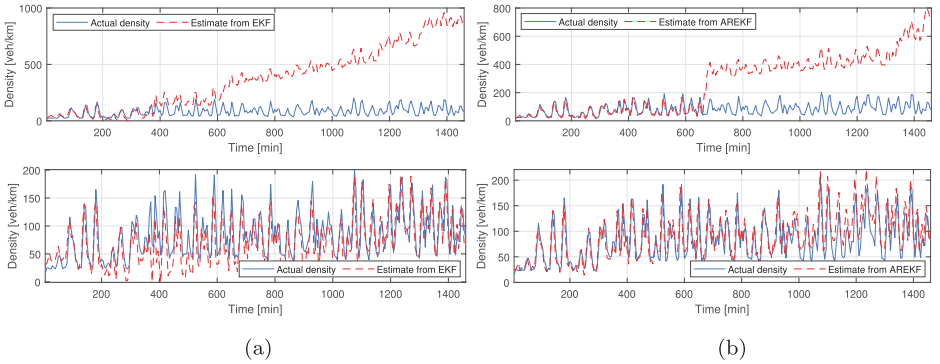

Fig. 10. Comparison of actual traffic density and estimated density from the (a) the EKF and (b) the AREKF with exit road missing data but without data imputation (top) and missing data imputation (bottom).

missing data (without imputation) results in the most inaccurate estimate, whereas the missing data imputation reduces the system sensitivity to missing data and the AREKF estimates the density accurately. However, the performance of AREKF with missing data is lower than with normal data, as expected, since missing data imputation is just an approximation to the actual system input.

The AREKF algorithm for traffic density estimation with normal data and missing data imputation of entry, merge, and exit flows takes about 400–500 milliseconds to compute, making it ideal for real-time implementation. Since the algorithm with missing data imputation has an additional processing stage, the computational time of the algorithm with missing data imputation is marginally longer.

3.3. Performance of adaptive ramp metering

The estimated traffic density from the AREKF is used to control the on-ramp traffic flow $q_s(k)$. The parameter λ_r is set at 70 veh/h based on real-world experiments. The performance of ALINEA (adaptive ramp) based on the estimated traffic density $\hat{\rho}(k)$ is evaluated for different traffic flow conditions (e.g., free flow to congested flow of a day) and compared the results to the fixed time ramp metering or fixed ramp (that utilizes historical traffic data and a time-of-day basis). We used various performance metrics to evaluate the comparison of simulation outcomes, e.g., traffic density, idling time, and traveling time as illustrated in Fig. 11(a). The results are plotted for every 10 minutes of data. We also show the distributions of performance parameters (Fig. 11(b)) for a better evaluation of our proposed method. Fig. 12 depicts how traffic density affects idling time and traveling time of vehicles. It is found that using the adaptive ALINEA, the traffic densities are clustered towards the lower density and can be kept close to the critical density at which the traffic is smoothest and the flow is highest.

The merging segment at this point allows the maximum number of vehicles to use the road without slowing it down, which reduces traveling time and idling time of vehicles. Moreover, it is observed that during the peak-hour, the fixed ramp metering causes the density to spike in between times 1000 to 1200 min, whereas for ALINEA the increase was insignificant. Thus, the ALINEA is able to prevent the shockwaves by controlling a sudden increase in density. The overall traffic flow performance by the proposed method is given in Table 2.

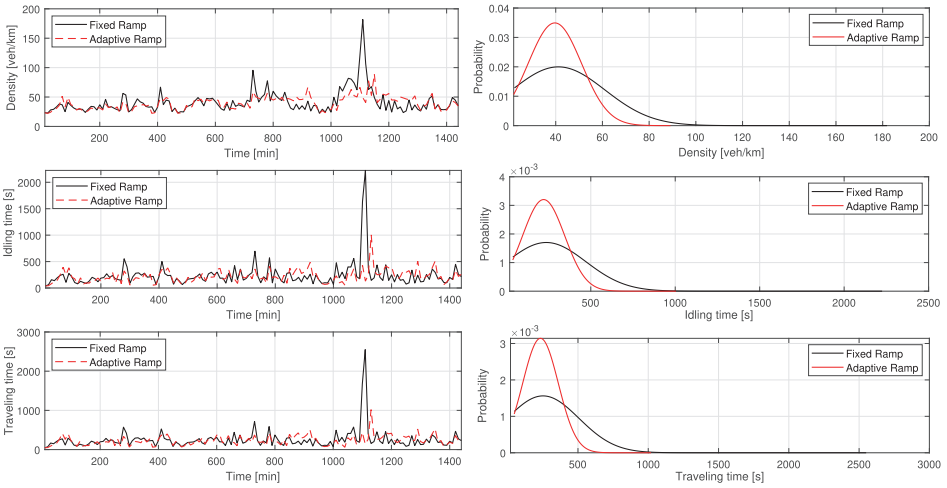

Fig. 11. Comparison of (a) performance metrics and (b) distributions of AREKF-based adaptive ramp metering proposed in this paper and fixed ramp metering for different traffic conditions of a day.

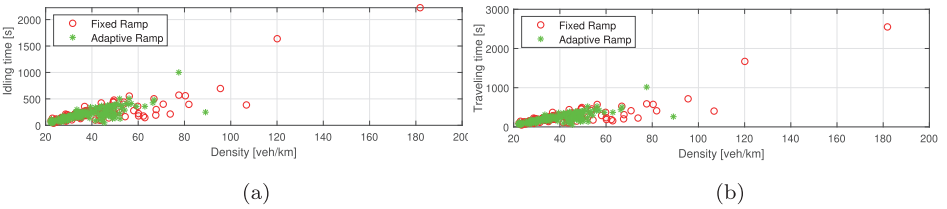

Fig. 12. (a) Idling time and (b) vehicle travel time in relation to traffic density. The proposed method groups traffic densities into lower-density clusters, which improves traffic performance in the merging road.

Table 2
Comparison of traffic flow performance with fixed and adaptive ramp metering .

Parameter	Fixed ramp	Adaptive ramp	% Improvement
Average density [Veh/km]	41.58	39.24	5.6
Average idling time [s]	236.48	220.56	6.7
Average traveling time [s]	251.64	233.28	7.3

4. Conclusion

In this paper, an AREKF has been developed for real-time traffic density estimation with missing data. We have applied a model-based missing data imputation technique to substitute missing data with estimated values as the pre-processing step of AREKF for reducing system input errors. A real-time ramp metering control algorithm called ALINEA is applied to adaptively control the flow rate of vehicles based on the estimated traffic density. The results show that the proposed AREKF can estimate the traffic density with sufficient accuracy if missing data occurs. Moreover, the ramp metering control has significantly improved the traffic flow performance by reducing traffic congestion and queue length. The traffic model proposed in this work is responsive to the system input and hence, the performance can be enhanced by modifying the model or decreasing the number of system inputs.

For future work, we could also apply correlation methods for time-correlated residual sequence due to sub-optimal estimation of the EKF. In addition, we could consider applying Kalman filters with simultaneous estimation of both the noise covariances in our traffic system.

References

[1] INRIX, Global traffic scorecard, (2019), Retrieved from https://inrix.com/press-releases/2019-traffic-scorecard-us/.
[2] A.R. Gner, A. Murat, R.B. Chinnam, Dynamic routing under recurrent and non-recurrent congestion using real-time ITS information, Computers & Operations Research 39 (2) (2012) 358–373.
[3] P.J. Jin, J. Fang, X. Jiang, M. DeGaspari, C.M. Walton, Gap metering for active traffic control at freeway merging sections, Journal of Intelligent Transportation Systems 21 (1) (2017) 1–11.

[4] F.A. Wagner, A.D. May Jr, Use of aerial photography in freeway traffic operations studies, Highway Research Record 19 (1963) 24–34.
[5] T.Z. Qiu, X.Y. Lu, A.H. Chow, S.E. Shladover, Estimation of freeway traffic density with loop detector and probe vehicle data, Transp Res Rec 2178 (1) (2010) 21–29.
[6] L. Muoz, X. Sun, R. Horowitz, L. Alvarez, Traffic density estimation with the cell transmission model, In Proc. IEEE American Control Conference 5 (2003) 3750–3755.
[7] L. Garelli, C. Casetti, C.F. Chiasserini, M. Fiore, Mobsampling: V2v communications for traffic density estimation, In Proc. IEEE 73rd Vehicular Technology Conference (2011) 1–5.
[8] B. Coifman, Estimating density and lane inflow on a freeway segment, Transportation Research Part A: Policy and Practice 37 (8) (2003) 689–701.
[9] L. Alvarez-Icaza, L. Munoz, X. Sun, R. Horowitz, Adaptive observer for traffic density estimation, In Proc. IEEE American Control Conference 3 (2004) 2705–2710.
[10] Y. Wang, M. Papageorgiou, Real-time freeway traffic state estimation based on extended kalman filter: a general approach, Transportation Research Part B: Methodological 39 (2) (2005) 141–167.
[11] N.E.E. Faouzi, H. Leung, A. Kurian, Data fusion in intelligent transportation systems: progress and challengesa survey, Information Fusion 12 (1) (2011) 4–10.
[12] K. Singh, B. Li, Estimation of traffic densities for multilane roadways using a markov model approach, IEEE Trans. Ind. Electron. 59 (11) (2011) 4369–4376.
[13] H. Andre, J.L. Ny, A differentially private ensemble kalman filter for road traffic estimation, in 2017 IEEE international conference on acoustics, Speech and Signal Processing (2017) 6409–6413.
[14] X. Di, H.X. Liu, G.A. Davis, Hybrid extended kalman filtering approach for traffic density estimation along signalized arterials: use of global positioning system data, Transp Res Rec 2188 (1) (2010) 165–173.
[15] H. Yang, P.J. Jin, B. Ran, D. Yang, Z. Duan, L. He, Freeway traffic state estimation: a Lagrangian-space Kalman filter approach, Journal of Intelligent Transportation Systems 23 (6) (2019) 525–540.
[16] M. St-Pierre, D. Gingras, Comparison between the unscented Kalman filter and the extended Kalman filter for the position estimation module of an integrated navigation information system, In IEEE Intelligent Vehicles Symposium (2004) 831–835.
[17] A.H. Mohamed, K.P. Schwarz, Adaptive kalman filtering for INS/GPS, J Geod 73 (4) (1999) 193–203.
[18] D. Lunni, G. Giordano, E. Sinibaldi, M. Cianchetti, B. Mazzolai, Shape estimation based on kalman filtering: towards fully soft proprioception, In IEEE International Conference on Soft Robotics (RoboSoft) (2018) 541–546.
[19] F. Sun, X. Hu, Y. Zou, S. Li, Adaptive unscented kalman filtering for state of charge estimation of a lithium-ion battery for electric vehicles, Energy 36 (5) (2011) 3531–3540.
[20] M. Behrisch, L. Bieker, J. Erdmann, D. Krajzewicz, SUMOSimulation of urban mobility: an overview, In Proc. 3rd International Conference on Advances in System Simulation (2011).
[21] T. Ajitha, L. Vanajakshi, S.C. Subramanian, Real-time traffic density estimation without reliable side road data, J. Comput. Civil Eng. 29 (2) (2015) 04014033.
[22] M. Grewal, A. Andrews, Kalman filtering: theory and practice using MATLAB, John Wiley & Sons (2001).
[23] D. Simon, Optimal state estimation: Kalman, H infinity, and nonlinear approaches, John Wiley & Sons, 2006.
[24] P.S. Maybeck, Stochastic models, estimation, and control, chapter 10 parameter uncertainties and adaptive estimation, 2, Academic Press, 1982, pp. 68–158.
[25] K. Reif, S. Gunther, E. Yaz, R. Unbehauen, Stochastic stability of the discrete-time extended Kalman filter, IEEE Trans Automat Contr 44 (4) (1999) 714–728.
[26] R.A. Johnson, D.W. Wichern, Applied multivariate statistical analysis, 6, Pearson, London, UK:, 2014.
[27] X. Xiao, K. Shen, Y. Liang, T. Liu, Kalman filter with recursive covariance estimation for protection against system uncertainty, IET Control Theory Appl. 14 (15) (2020) 2097–2105.
[28] S. Akhlaghi, N. Zhou, Z. Huang, Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation, IEEE power & energy society general meeting (2017) 1–5.
[29] W. Gao, J. Li, G. Zhou, Q. Li, Adaptive kalman filtering with recursive noise estimator for integrated SINS/DVL systems, J. Navig. 68 (1) (2015) 142–161.
[30] A. Almagbile, J. Wang, W. Wang, Evaluating the performances of adaptive Kalman filter methods in GPS/INS integration, Journal of Global Positioning Systems 9 (1) (2010) 33–40.
[31] R. Mehra, On the identification of variances and adaptive kalman filtering, IEEE Trans Automat Contr 15 (2) (1970) 175–184.
[32] R. Mehra, Approaches to adaptive filtering, IEEE Trans Automat Contr 17 (5) (1972) 693–698.
[33] X.Y. Lu, A. Skabardonis, Freeway traffic shockwave analysis: exploring the NGSIM trajectory data, In 86th Annual Meeting of the Transportation Research Board, 2007. Washington, DC
[34] I. Papamichail, M. Papageorgiou, Traffic-responsive linked ramp-metering control, IEEE Trans. Intell. Transp. Syst. 9 (1) (2008) 111–121.
[35] C. Demiral, H.B. Celikoglu, Application of ALINEA ramp control algorithm to freeway traffic flow on approaches to Bosphorus strait crossing bridges Procedia-Social and Behavioral Sciences 20 (2011) 364–371.
[36] M.J. Cassidy, K. Jang, C.F. Daganzo, Macroscopic fundamental diagrams for freeway networks: theory and observation, Transp Res Rec 2260 (1) (2011) 8–15.
[37] C.D. Lewis, Industrial and business forecasting methods: apractical guide to exponential smoothing and curve fitting, Butterworth-Heinemann (1982).