Simultaneous Traffic Flow and Macro Model Estimation for Signalized Junctions with Multiple Input Lanes

Luana Chetcuti Zammit, Simon G. Fabri and Kenneth Scerri
Department of Systems and Control Engineering, University of Malta, Msida, Malta MSD2080, Malta

Keywords: Macro Model Estimation, Expectation-maximization Algorithm, Quasi Real-time Estimation, Traffic Junction Modelling, Urban Traffic Control, Intelligent Transportation System.

Abstract: A novel algorithm is presented for macro model estimation of the dynamics of traffic flow in a junction having multiple input lanes for each turning direction. The proposed algorithm jointly estimates the states describing the traffic flow under different traffic conditions, together with model parameters and their uncertainties of the measurement and process noise. Use is made of the Expectation-Maximization methodology with a sliding window over time in order to obtain quasi real-time estimation.

1 INTRODUCTION

Automatic control of signalized traffic intersections is a vital component of modern traffic infrastructures. The aim is to optimize the flow of vehicles through the road network in the presence of time-varying traffic conditions. Thus, the design and implementation of such systems requires the use of computationally efficient numerical models that continuously estimate the dynamics of traffic flow.

There are two main classes of traffic modelling techniques - *macroscopic* or *microscopic*. Macroscopic models estimate traffic quantities at a high level of aggregation (Lighthill and Whitham, 1955; Richards, 1956), whereas microscopic models capture the dynamic behaviour of individual vehicles. For the purpose of automated traffic light control, macroscopic models are stronger candidates due to a lower computation demand and faster real-time estimation. Hence, macroscopic models are vital to the development of controllers capable to adapt to changing traffic behaviour.

Hence, this study is based on a data-driven, macroscopic traffic junction model. This model is rewritten in state-space form to allow the adoption of various versatile control techniques from systems theory. The model makes use of queuing theory to describe the traffic dynamics. Queuing theory has been applied to transport in classical works, (Beckmann et al., 1955; Webster, 1957) and more recently in the works of Olszewski (1994); Viti and van Zuylen (2004). Moreover, macro models in state-space form making use of queuing theory to describe the dynamics of traffic flow in a junction with a single input lane have been presented in Homolova (2005); Kratochvilova and Nagy (2004); Pecherkova et al. (2008); Zammit et al. (2016). In this study this approach is extended to describe the dynamics of traffic flow in a junction having multiple input lanes for each turning direction.

Additionally, traffic quantities are estimated by the model in real-time, based only upon a few basic sensor measurements and assuming little knowledge of the underlying traffic parameters. The number of sensors in a junction is kept to a minimum so as to reduce infrastructural costs, as discussed in more detail in the next section.

This leads to a novel algorithm to estimate jointly, in quasi real-time, the model states, the unknown and possibly time-varying model parameters and noise covariances. This differs from the implementations of previous works (Homolova, 2005; Kratochvilova and Nagy, 2004; Pecherkova et al., 2008) where only the estimation of model states was carried out and the model parameters were tuned a priori from past traffic measurements or from simulated data using software such as Aimsun, SUMO and VISUM. Hence no real-time updates of the parameters were obtained from the previous works.

The joint estimation is based on the Expectation-Maximization (EM) algorithm (Dempster et al., 1977), but modified to obtain quasi real-time implementation by utilising a sliding window (Dang et al., 2009). Despite the successful application of EM

Zammit, L., Fabri, S. and Scerri, K.
Simultaneous Traffic Flow and Macro Model Estimation for Signalized Junctions with Multiple Input Lanes.
DOI: 10.5220/0006244801570164
In *Proceedings of the 3rd International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2017)*, pages 157-164
ISBN: 978-989-758-242-4

methods to several other fields, their potential utility to traffic flow models as explored in this work is an innovative contribution.

2 MODEL DEVELOPMENT

This work assumes that for a junction with two inflow directions, only three sensors are installed per arm, usually implemented as inductive loops (Dunn Engineering Associates and Siemens Intelligent Transporation Systems, 2005). Two are placed at the input lane to measure the inflow towards the junction for each turning direction. The other sensor is placed next to the output lane of the arm to measure the outflow away from the junction as shown in Figure 1.

The variables denoting traffic flow through the junction are: $\gamma_I(t)$ which represents the number of unit vehicles (uv) entering an arm in a cycle in [uv/cycle] where a cycle is the time required for one complete sequence of traffic signal phases; lane occupancy $\phi(t)$ which is the proportion of time when a sensor is occupied (and therefore activated) in a cycle with respect to the total measuring period given in [%]; and $\zeta(t)$ which represents the number of cars waiting to pass through the intersection at the start of the red phase of each cycle (in $[uv]$).

The sensor measurements include i) $\gamma_I(t)$ in [uv/cycle], ii) $\phi(t)$ and iii) the outflow from an arm, denoted as $\gamma_O(t)$ in [uv/cycle], which represents the number of unit vehicles exiting an arm during the green signal in a given cycle.

Let $\gamma_I(t)$, $\phi(t)$ and $\zeta(t)$ be the state variables of the model's state space equations with integer t denoting the cycle index.

Traffic flow performance is subject to the so-called unsaturated or saturated flow conditions (Gazis, 2002). Saturated flow corresponds to the maximal number of vehicles that could flow through an arm. This depends on characteristics such as road width, number of traffic lanes in one direction, turning movements and speed limits. Assume this maximal flow to be a positive constant S [uv/cycle]. Otherwise the arm is unsaturated.

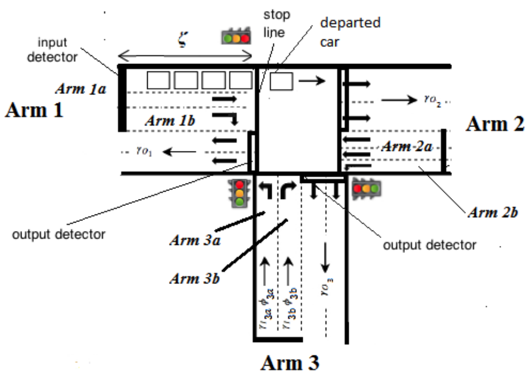

Figure 1: 3-arm signalized junction with two input lanes for each arm representing different turning directions.

The junction dynamics for a given arm can be represented using the following notation. Let z represent the ratio of the green signal time for an arm to the total cycle time. $\mathbf{w}(\mathbf{t})$ is a white, zero-mean, Gaussian noise process with covariance $\mathbf{Q}$, capturing model inaccuracy. $\mathbf{Q}$ is a diagonal matrix with $q_{l,j}$ representing the elements of the inverse of $\mathbf{Q}$ and l and j represent the rows and columns of the matrix respectively. $\mathbf{v}(\mathbf{t})$ is a white, zero-mean, Gaussian measurement noise with covariance $\mathbf{R}$, capturing sensor deficiencies. $\mathbf{R}$ is a diagonal matrix with $r_{l,j}$ representing the elements of the inverse of $\mathbf{R}$. A binary queue length indicator $\delta(t)$ is defined that takes a value of 0 under unsaturated conditions and 1 if a lane is saturated.

The queue length, $\zeta(t+1)$ is described by a piecewise linear throughput characteristic at the intersection, which applies the principle of conservation of traffic flow (Pecherkova et al., 2008). If $I_O(t)$ represents the number of unit vehicles exiting an arm in a cycle if a continuous green signal is shown throughout the cycle, then the number of vehicles exiting during a cycle with green ratio $z(t)$, denoted as $\gamma_O(t)$, is given by $z(t)I_O(t)$. Thus it follows that $\zeta(t+1)$ depends on the previous queue length, $\zeta(t)$, the departed vehicles, $\gamma_O(t)$ and the arrived cars $\gamma_I(t)$ in a cycle, as given by Equation (1):

$$\zeta(t+1) = \zeta(t) - \gamma_O(t) + \gamma_I(t) \tag{1}$$

For unsaturated traffic conditions, the model assumes that the outflow is equal to the inflow if no queue $\zeta(t)$ exists, otherwise the outflow increases according to the queue $\zeta(t)$ and green time z. Hence for unsaturated traffic conditions, $(\gamma_O(t) < S(t))$ the following holds:

$$I_O(t) = \gamma_I(t) + \frac{\zeta(t)}{z(t)}$$

$$\implies \gamma_O(t) = z(t)\gamma_I(t) + \zeta(t) \tag{2}$$

For saturated traffic conditions, $I_O(t)$ is equal to $S(t)$ which implies that:

$$\gamma_O(t) = z(t)S(t) \tag{3}$$

Substituting for $\gamma_O(t)$ in Equation (1) for both unsaturated and saturated traffic conditions, the queue length at sample instant $t+1$ is given by Equations (4) and (5) respectively.

$$\zeta(t+1) = (1 - z(t))\gamma_I(t) \tag{4}$$

$$\zeta(t+1) = \zeta(t) + \gamma_I(t) - z(t)S(t) \tag{5}$$

A Markovian random process is assumed to model the dynamics of the inflow to the junction, γ_I. Also the occupancy $\phi(t+1)$ is considered to depend upon the occupancy at the previous cycle $\phi(t)$ and the queue length $\zeta(t)$, linearly parameterized by two variables $\kappa(t)$ and $\beta(t)$. The occupancy is a useful measurement and is included to detect unusual situations, such as road blockage during unsaturated traffic conditions.

The above equations can be represented in discrete-time stochastic state-space form (6) and (7) where the sampling period is taken to be the cycle time. Unsaturated traffic conditions are represented by Equation (6) and saturated traffic conditions by Equation (7). The second equation in both (6) and (7) represents the sensor readings of the controlled intersection. The model equations (6) and (7) are limited to one arm only, but they can be expanded to model multiple arms in an intersection, indexed by $i = 1, 2, ..., n$. In a 3-arm junction with bi-directional traffic flow and separate lanes for different turning directions towards arms a or b respectively as shown in Figure 1, the output flow from each arm and hence the number of cars leaving the junction, are represented as $\gamma_{O_1}(t)$, $\gamma_{O_2}(t)$ and $\gamma_{O_3}(t)$ respectively. The outflow for each arm, for unsaturated traffic conditions and for saturated traffic conditions are respectively given by:

$$\begin{aligned}
\gamma_{O_1}(t) &= (\zeta_{2a}(t) + z_{2a}\gamma_{I_{2a}}(t)) + (\zeta_{3a}(t) + z_{3a}\gamma_{I_{3a}}(t)) \\
\gamma_{O_2}(t) &= (\zeta_{1a}(t) + z_{1a}\gamma_{I_{1a}}(t)) + (\zeta_{3b}(t) + z_{3b}\gamma_{I_{3b}}(t)) \\
\gamma_{O_3}(t) &= (\zeta_{1b}(t) + z_{1b}\gamma_{I_{1b}}(t)) + (\zeta_{2b}(t) + z_{2b}\gamma_{I_{2b}}(t))
\end{aligned}$$

$$\begin{aligned}
\gamma_{O_1}(t) &= S_{2a} \cdot z_{2a} + S_{3a} \cdot z_{3a} \\
\gamma_{O_2}(t) &= S_{1a} \cdot z_{1a} + S_{3b} \cdot z_{3b} \\
\gamma_{O_3}(t) &= S_{1b} \cdot z_{1b} + S_{2b} \cdot z_{2b}
\end{aligned}$$

where S_{ia} represents the saturation flow for arm ia and S_{ib} represents the saturation flow for arm ib where i=1, 2, 3.

3 JOINT ESTIMATION OF STATES, PARAMETERS AND NOISE

Joint estimation refers to the process of simultaneously estimating the state of a dynamic system and the model which gives rise to the dynamics, including all model parameters and covariances of the process and measurement noise.

Published works on joint estimation methods make use of nonlinear estimation algorithms such as the extended Kalman filter (Wang et al., 2008), or particle filtering (Mihaylova et al., 2007). However, divergence problems associated with such joint estimation methods and the high computational demands associated with particle filters can restrain their applications (Huber, 2015). Furthermore, the Maximum Likelihood (ML) principle (Ljung, 1999) plays a key role in joint estimation. ML estimation, solved via a gradient-based search strategy such as a Newton type method (Soderstrom and Stoica, 1989) is sometimes difficult to solve for state-space models. Hence, a gradient-search free computation of the ML, such as the EM algorithm (Dempster et al., 1977) is applied in this work to infer both the state space model states and its parameters.

For a n-arm junction, the unknown variables to be estimated include the state vector $\mathbf{x} \triangleq [\zeta_{1a}, \zeta_{1b} \, ... \, \zeta_{na} \, \zeta_{nb} \, \gamma_{I_{1a}} \, \gamma_{I_{1b}} \cdots \gamma_{I_{na}} \, \gamma_{I_{nb}} \, \phi_{1a} \, \phi_{1b} \, ... \, \phi_{na} \, \phi_{nb}]^T$, the vector of model parameters $\boldsymbol{\theta} \triangleq [\, \kappa_{1a} \, \kappa_{1b} \, ... \, \kappa_{na} \, \kappa_{nb} \, \beta_{1a} \, \beta_{1b} \, ... \, \beta_{na} \, \beta_{nb} \, S_{1a} \, S_{1b} \, ... \, S_{na} \, S_{nb}]^T$ and noise covariances $\mathbf{Q}$, $\mathbf{R}$. This work extends that in Zammit et al. (2016), where instead of noise realisations, the measurement and process noise covariances are estimated. Furthermore, the model parameters $\boldsymbol{\theta}$ are considered separate from the noise covariances $\mathbf{Q}$ and $\mathbf{R}$, as reflected in the estimation algorithm of Section 3.2. This differs from the standard EM where $\mathbf{Q}$ and $\mathbf{R}$ are typically grouped with the model parameters (Bishop, 2009). Our approach thus allows the estimation algorithm to be tuned according to the differing characteristics of the model parameters and the elements of the covariance matrices, such as their different orders of magnitude and their numerical constraints.

Futhermore, in this study the classical EM algorithm is modified to effect quasi real-time estimation since the standard EM is a multiple pass batch processing algorithm, where estimation is carried out offline based on a batch of measurements available a priori in time as described in the next section.

3.1 Standard Batch-based EM

Let $\hat{\mathbf{x}}$ denote the estimate of the state vector $\mathbf{x}$, where $\hat{\mathbf{x}} \triangleq [\hat{\zeta}_{1a}, \hat{\zeta}_{1b} \, ... \, \hat{\zeta}_{na} \, \hat{\zeta}_{nb} \, \hat{\gamma}_{I_{1a}} \, \hat{\gamma}_{I_{1b}} \cdots \hat{\gamma}_{I_{na}} \, \hat{\gamma}_{I_{nb}} \, \hat{\phi}_{1a} \, \hat{\phi}_{1b} \, ... \, \hat{\phi}_{na} \, \hat{\phi}_{nb}]^T$. In the standard EM, $\hat{\mathbf{x}}$ is given by running the Kalman Smoother recursions (Sarkka, 2013). To solve for $\boldsymbol{\theta}$, the parameters that maximize the below objective function (Chen, 2006) conditioned upon the estimated states, are iteratively estimated over a batch of N observations,

$$G(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = E(log P(\mathbf{x}, \mathbf{y}|\boldsymbol{\theta})|\mathbf{y}, \hat{\boldsymbol{\theta}})$$

where $\hat{\boldsymbol{\theta}}$ is an estimate of $\boldsymbol{\theta}$ and defined as $\hat{\boldsymbol{\theta}} \triangleq [\, \hat{\kappa}_{1a} \, \hat{\kappa}_{1b} \, ... \, \hat{\kappa}_{na} \, \hat{\kappa}_{nb} \, \hat{\beta}_{1a} \, \hat{\beta}_{1b} \, ... \, \hat{\beta}_{na} \, \hat{\beta}_{nb} \, \hat{S}_{1a} \, \hat{S}_{1b} \, ... \, \hat{S}_{na} \, \hat{S}_{nb}]^T$.

Hence, on the k^{th} iteration, partial differentiation of the objective function with respect to each of the elements of $\hat{\boldsymbol{\theta}}$ is performed, set to zero and solved simultaneously. Due to the large number of parameters to be estimated, the equations are solved using

Unsaturation Case:

$$\begin{bmatrix} \zeta(t+1) \\ \gamma_I(t+1) \\ \phi(t+1) \end{bmatrix} = \begin{bmatrix} 0 & 1-z(t) & 0 \\ 0 & 1 & 0 \\ \kappa(t) & 0 & \beta(t) \end{bmatrix} \begin{bmatrix} \zeta(t) \\ \gamma_I(t) \\ \phi(t) \end{bmatrix} + \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{bmatrix}$$
$$\begin{bmatrix} \gamma_I(t) \\ \phi(t) \\ \gamma_O(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & z(t) & 0 \end{bmatrix} \begin{bmatrix} \zeta(t) \\ \gamma_I(t) \\ \phi(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \end{bmatrix} \quad (6)$$

Saturation Case:

$$\begin{bmatrix} \zeta(t+1) \\ \gamma_I(t+1) \\ \phi(t+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ \kappa(t) & 0 & \beta(t) \end{bmatrix} \begin{bmatrix} \zeta(t) \\ \gamma_I(t) \\ \phi(t) \end{bmatrix} - \begin{bmatrix} S(t) \\ 0 \\ 0 \end{bmatrix} z(t) + \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{bmatrix}$$
$$\begin{bmatrix} \gamma_I(t) \\ \phi(t) \\ \gamma_O(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \zeta(t) \\ \gamma_I(t) \\ \phi(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ S(t) \end{bmatrix} z(t) + \begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \end{bmatrix} \quad (7)$$

least squares estimation with positive constraints on the saturation parameters.

Taking the partial derivative of the objective function with respect to κ_{1a} and equating to 0, gives:

$$\hat{\kappa}_{1a}(k)\sum_{k=t}^{N}\hat{q}_{11,11}(k)\hat{\zeta}_1a(k-1)^2 = \sum_{k=t}^{N}\hat{q}_{11,11}(k)\hat{\phi}_{1a}(k)\hat{\zeta}_{1a}(k-1)\} - \hat{\beta}_{1a}(k)\sum_{k=t}^{N}\hat{q}_{11,11}(k)\hat{\phi}_{1a}(k-1)\hat{\zeta}_{1a}(k-1) \quad (8)$$

Similarly for $\hat{\beta}_{1a}$:

$$\hat{\beta}_{1a}(k)\sum_{k=t}^{N}\hat{q}_{11,11}(k)\hat{\phi}_{1a}(k-1)^2 = \sum_{k=t}^{N}\{\hat{q}_{11,11}(k)\hat{\phi}_{1a}(k)\hat{\phi}_{1a}(k-1)\} - \hat{\kappa}_{1a}(k)\sum_{k=t}^{N}\hat{q}_{11,11}(k)\hat{\phi}_{1a}(k-1)\hat{\zeta}_{1a}(k-1) \quad (9)$$

For $\hat{S}_{1a}$ under saturated traffic conditions:

$$\sum_{k=t}^{N}\Big(\hat{q}_{1,1}(k)[\hat{\zeta}_{1a}(k-1)+\hat{\gamma}_{I_{1a}}(k-1)]z_{1a}(k-1) - \hat{q}_{1,1}(k)\hat{\zeta}_{1a}(k)z_{1a}(k-1) + \hat{r}_{15,15}(k)\hat{\gamma}_{O_3}(k)z_{1a}(k) - \hat{r}_{15,15}(k)\hat{S}_{2b}(k)z_{2b}(k)z_{1a}(k)\Big) = \hat{S}_{1a}(k)\sum_{k=t}^{N}\hat{r}_{15,15}(k)\Big(z_{1a}(k-1)^2+z_{1a}(k)^2\Big) \quad (10)$$

Similarly, the covariances $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$ of the noise are estimated by Equations (11) and (12), where $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$ and $\hat{\mathbf{D}}$ are the estimated state space matrices and $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$ are the estimated covariances of $\mathbf{Q}$ and $\mathbf{R}$ with $\hat{r}_{l,j}$ and $\hat{q}_{l,j}$ being the estimates of $r_{l,j}$ and $q_{l,j}$ respectively. The quantities $\mathbf{P}_{k|N}$, $\mathbf{P}_{k-1,k}$, $\mathbf{P}_{k,k-1}$ and $\mathbf{P}_{k-1|N}$ are pre-computed from the Kalman Smoother recursions (Chen, 2006).

$$\hat{\mathbf{Q}} = \frac{1}{(N-1)}\sum_{k=2}^{N}\Big[(\hat{\mathbf{x}}_{k|N}-\hat{\mathbf{A}}(k-1)\hat{\mathbf{x}}_{k-1|N}-\hat{\mathbf{B}}(k-1)\mathbf{z}(k-1)) (\hat{\mathbf{x}}_{k|N}-\hat{\mathbf{A}}(k-1)\hat{\mathbf{x}}_{k-1|N}-\hat{\mathbf{B}}(k-1)\mathbf{z}(k-1))^T + \mathbf{P}_{k|N}-\hat{\mathbf{A}}(k-1)\mathbf{P}_{k-1,k}-\mathbf{P}_{k,k-1}\hat{\mathbf{A}}(k-1)^T + \hat{\mathbf{A}}(k-1)\mathbf{P}_{k-1|N}\hat{\mathbf{A}}(k-1)^T\Big] \quad (11)$$

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{k=1}^{N}\Big[(\mathbf{y}(k)-\hat{\mathbf{C}}(k-1)\hat{\mathbf{x}}_{k|N}-\hat{\mathbf{D}}(k-1)\mathbf{z}(k)) (\mathbf{y}(k)-\hat{\mathbf{C}}(k-1)\hat{\mathbf{x}}_{k|N}-\hat{\mathbf{D}}(k-1)\mathbf{z}(k))^T + \hat{\mathbf{C}}(k-1)\mathbf{P}_{k|N}\hat{\mathbf{C}}(k-1)^T\Big] \quad (12)$$

This process is repeated for all other variables describing the traffic dynamics within a signalized junction for unsaturated and saturated traffic conditions. Equations (8) to (12) denote a batch algorithm which is not suitable for real-time estimation and control. Hence a novel modified algorithm for joint estimation of states, noise and model parameters in quasi real-time is proposed in the following section.

3.2 The Modified EM Algorithm

The modified EM algorithm is represented in Table 1. Three main considerations are tackled: i) quasi real-time implementation, ii) switching conditions and iii) probing.

i) Quasi Real-Time Implementation: In practical dynamic traffic situations, traffic conditions change in real-time. Hence, the standard EM algorithm requiring a sizeable batch of N data points is not suitable for online estimation. Instead, to carry out quasi real-time estimation, the iterative algorithm in Table 1 makes use of two uniform windows, one of fixed time length $\bar{n}$ for joint states and parameter estimation, and another of fixed time length $\bar{m}$ for noise covariance estimation, hence separating the estimation of $\mathbf{Q}$ and $\mathbf{R}$ from the estimation of model parameters, as already discussed. To obtain optimal estimation results, different window sizes, $\bar{n}$ and $\bar{m}$ are allowed, but both are significantly less than N to obtain quasi real-time estimation. Joint parameter and state estimation is carried out for those particular time points falling inside the first window, by applying the EM algorithm presented in Table 1. At every time iteration, the window slides forward by one instant and the procedure is repeated again, with initial values for the parameters and the states being fed from the previous time window. In the second time window, noise covariance estimation is carried out for those particular time points falling inside this window. Tests to establish suitable window lengths $\bar{m}$ and $\bar{n}$ were carried out through simulations as described in Section IV.

Table 1: EM algorithm for estimation of model parameters.

Initialise estimates for $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$
Commencing from $t{=}\bar{n}+1$
Iterate for every time step t and measure $\mathbf{y}(t)$.
  Iterate for $k{=}(t-\bar{n}),...,t$

  E-step
    Run Kalman-Filter recursions followed by the Kalman Smoother recursions in order to compute $\hat{\mathbf{x}}_{k|\bar{n}}$.

  M-step
    Maximise $G(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_k)$ over $\boldsymbol{\theta}$ for unsaturated conditions including Equations (8)-(9) and for saturated conditions including Equations (8)-(10) with N replaced by $\bar{n}$.
    Repeat until the log likelihood of the objective function converges to a constant value up to a small predefined tolerance bound.
    Update $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$, with $\hat{\boldsymbol{\theta}}_k$ to reflect the traffic conditions per arm.
    Example: $\hat{\mathbf{A}}{=}\frac{(\hat{\mathbf{A}}_{unsat}\cdot\bar{n}_{unsat})(\hat{\mathbf{A}}_{sat}\cdot\bar{n}_{sat})}{\bar{n}}$
    where $\hat{\mathbf{A}}_{unsat}$ and $\hat{\mathbf{A}}_{sat}$ represent the state transition matrix for unsaturation and saturation respectively, while $\bar{n}_{unsat}$ and $\bar{n}_{sat}$ represent the number of times an arm is unsaturated or saturated. Hence $\bar{n}{=}\bar{n}_{unsat}{+}\bar{n}_{sat}$.

  If $(t{>}\bar{m})$ where m represents the size of the second window
    Iterate for $k{=}(t-\bar{m}),...,t$
    Maximise $\mathbf{Q}$ and $\mathbf{R}$ as in Equations (11)-(12) with N replaced by $\bar{m}$.

  t=t+1

This algorithm reduces the computational and storage demands for a junction since traffic information from sensors, including the inflow and the outflow, need not be available as a batch of N samples, but rather within a much shorter window of $\bar{n}$ or $\bar{m}$ samples.

ii) Switching Conditions: The presented model is subject to changing traffic conditions, i.e. unsaturated or saturated traffic conditions per arm and direction. An arm can exhibit unsaturated or saturated behaviour irrespective of other arms within the same junction. An arm is saturated if the condition $(\hat{\zeta}_{nd}(k)+z_{nd}(k)\hat{\gamma}_{I_{nd}}(k)){\geq}(\hat{S}_{nd}(k-1)z_{nd}(k))$ is satisfied with $d{=}a$ or b denoting the arm for each turning direction. Conversely, an arm is unsaturated if $(\hat{\zeta}_{nd}(k)+z_{nd}(k)\hat{\gamma}_{I_{nd}}(k)){<}(\hat{S}_{nd}(k-1)z_{nd}(k))$. These switching conditions are estimated through the modified EM algorithm results. For example, for a three arm junction with two separate lanes for each direction a and b, sixty-four ($2^{3\times2}$) different combinations of switching conditions exist. Thus during execution of the modified EM algorithm, the state space model's matrices are all updated to reflect the different traffic conditions per arm through the use of average weighting. For one time instance, falling within a window length $\bar{n}$, the number of times an arm is in unsaturation or saturation is noted and average weighting of model parameters and state space matrices are obtained as indicated in Table 1.

iii) Probing: A small probing dither signal, consisting of zero mean Gaussian noise with variance of 1×10^{-8} is introduced on the input green timing to make the parameter estimation process more efficient and to elicit richer information about the unknown parameters. This way, the system input is made more persistently exciting to encourage better estimation of the time-varying parameters (Astrom and Wittenmark, 1995).

4 RESULTS

The proposed algorithm for joint estimation of states, model parameters and noise covariances was tested and validated on a signalized 3-arm junction, with geometry similar to the junction presented in Figure 1. Two cases were tested: i) the estimation of states and model parameters, with noise covariances assumed known and ii) the estimation of states and model parameters together with process and measurement noise covariances.

A Root Mean Square Error (RMSE) measure is defined to determine the accuracy of the estimation results.

Definition 4.1. For some estimate $\hat{p}$, the RMSE is given by the square root of the averaged mean square error per sample in the time window, defined as:

$$J \triangleq \sqrt{\frac{\sum_{\bar{n}}(p(t)-\hat{p}(t))^2}{\bar{n}}} \tag{13}$$

where p is the actual value.

For the signalized 3-arm junction, due to the unavailability of actual data from a real junction, measurements of cars entering and leaving each arm were simulated in Aimsun as *Traffic State* per second. Traffic light information such as phases and a cycle time of 90 seconds were introduced to reflect typical traffic characteristics. The simulation was executed to generate traffic count measurements similar to a physical ITS junction purposely fitted with sensors.

Tests were first carried out to determine suitable window sizes for the modified EM algorithm. To determine $\bar{n}$, the expectation stage was executed separately from the maximization stage with different window sizes. In addition, Monte Carlo runs with 1000 different realisations were executed for the maximization stage, with different window size $\bar{n}$, for different traffic conditions. As expected it was noted that the accuracy of the estimation of states and parameters improved with increasing time lags. However, a balance between computation efficiency and estimation accuracy needs to be sought. For the inflow, under saturated traffic conditions, with $\bar{n} = 20$, the value of J obtained was 11.5% of the mean inflow measurements. With $\bar{n} = 40$, it was 8.5%, thus having only a 3% decrease in accuracy. Hence $\bar{n} = 20$ was preferred over 40.

To determine $\bar{m}$, Monte Carlo runs with 1000 different realisations were executed for the maximization stage of $\mathbf{Q}$ and $\mathbf{R}$. The matrix Euclidean norm of the estimated covariances, $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$, represented as $||\hat{\mathbf{Q}}||$ and $||\hat{\mathbf{R}}||$ were calculated for different window sizes and compared with the Euclidean norm of $\mathbf{Q}$ and $\mathbf{R}$ as shown in Table 2 and 3. With only a 0.02% difference in the Euclidean norm, $\bar{m} = 1500$ was preferred over 2000.

The training data generated by Aimsun, which includes $\boldsymbol{\gamma}_{I_i}$, $\boldsymbol{\phi}_i$, $\boldsymbol{\gamma}_{O_i}$, $i = 1,2,3$, was used to generate the sensor readings. The EM algorithm was then executed to jointly estimate the states, the model parameters and noise covariances. Table 4 shows some of the results obtained for the saturation values of this junction compared with the actual values, averaged over the whole training time, with a satisfactory % difference ranging from 0.121% to 1.025%.

To further test the model under different traffic conditions, fresh validation datasets were generated from Aimsun. The previously defined measure J for both known and unknown covariances is shown in Table 5. Here J is expressed as a percentage of $\sqrt{\frac{\sum_{\bar{n}} p^2(t)}{\bar{n}}}$ to yield a normalized measurement over one window. For comparison reasons, one figure of merit was computed for both tests shown in the last row of Table 5. By taking the resultant mean value over all the 12 individual % RMSE estimates for each test, an average of 0.501% was obtained when the noise covariances are known, whilst the average with estimation of noise covariances was 0.648%. Although the reduction in accuracy in the second case is very minor, such a reduction is expected since more variables were estimated than in the first case.

Figure 2 shows the results for γ_{O_1}, one arbitrarily selected parameter with noise covariance estimation. Its corresponding parameter as calculated from Aimsun is also superimposed showing that the model es-

Table 2: Estimated Covariance $\hat{\mathbf{Q}}$.

Window size $\bar{m}$	$\|\|\hat{\mathbf{Q}}\|\|$	$\|\|\mathbf{Q}\|\|$	% Difference
1000	4.302	1×10^{-7}	4.302×10^{9}
1500	9.998×10^{-8}	1×10^{-7}	-0.02
2000	1×10^{-7}	1×10^{-7}	0.00

Table 3: Estimated Covariance $\hat{\mathbf{R}}$.

Window size $\bar{m}$	$\|\|\hat{\mathbf{R}}\|\|$	$\|\|\mathbf{R}\|\|$	% Difference
1000	0.110	1×10^{-7}	1.10×10^{8}
1500	1×10^{-7}	1×10^{-7}	0.00
2000	1×10^{-7}	1×10^{-7}	0.00

Table 4: Estimated results for saturation parameters.

Saturation	Estimated Mean	Expected mean	% Difference
$\hat{S}_{1a}$	125.411	126	0.467
$\hat{S}_{1b}$	48.865	49	0.276
$\hat{S}_{2a}$	103.874	104	0.121
$\hat{S}_{2b}$	53.799	54	0.372
$\hat{S}_{3a}$	31.961	32	0.121
$\hat{S}_{3b}$	50.477	51	1.025

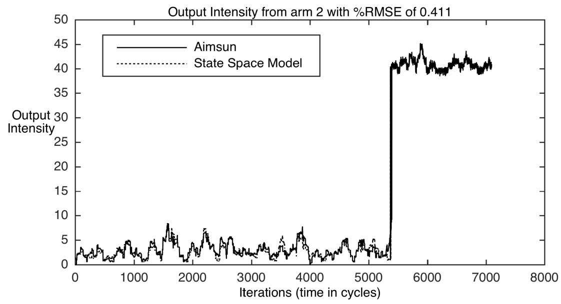

Figure 2: One Step Ahead Estimations of outflow.

timates compare highly with the ones obtained from Aimsun. Aimsun, being a microsimulator, implements a car-following model known as the Gipps model (Gipps, 1981) where vehicles accelerate to achieve the desired speed and decelerate when drivers have to avoid a collision, while trying to maintain the desired speed. On the other hand, this study provides a macrosocopic model, resulting in the general evaluation of the traffic flow behaviour, rather than individual driver behaviour with its acceleration and deceleration instances. Nevertheless, despite this different approach, Figure 2 only shows very minor discrepancies between the results, which could be attributed to the significant different nature of macro and micro models.

In Pecherkova et al. (2008), where known parameters are assumed, $J = 3.5897$ was obtained for the queue length, with data exhibiting a maximum value of 40. This represents a RMSE of 8.974% of the maximum. In our case, with joint estimation of parameters and noise covariances, for the queue lengths of each arm we obtained a RMSE of 8.840%, 5.473% and 6.892% of the maximum respectively, resulting in a relatively smaller average RMSE of 7.068%.

Table 5: Results on validation datasets with and without noise estimation.

Estimate	% RMSE with known noise covariance	% RMSE with noise covariance estimation
$\hat{\zeta}_1$	0.016	0.036
$\hat{\zeta}_2$	0.333	0.335
$\hat{\zeta}_3$	1.233	1.248
$\hat{\gamma}_{O_1}$	0.427	0.488
$\hat{\gamma}_{O_2}$	0.401	0.411
$\hat{\gamma}_{O_3}$	0.123	0.123
$\hat{\phi}_{1a}$	0.170	0.560
$\hat{\phi}_{1b}$	0.195	0.484
$\hat{\phi}_{2a}$	0.427	1.104
$\hat{\phi}_{2b}$	0.210	0.267
$\hat{\phi}_{3a}$	1.311	1.449
$\hat{\phi}_{3b}$	1.166	1.269
Average % RMSE	0.501	0.648

5 CONCLUSIONS

To contribute to the autonomicity of traffic light systems, this work proposes a quasi real-time macro model self-estimation method for the state variables, model parameters and noise covariances describing the dynamics of traffic flow in a junction with multiple lanes for each arm. Unlike previous works, the model parameters are not assumed to be known apriori. Modifications to the batch approach of the EM algorithm are presented to jointly estimate the states, the parameters and the noise covariances of the model in quasi real-time, by using small time windows of measurements. The results compare well with Pecherkova et al. (2008) where model parameters are assumed to be known and state estimation only is performed for much simpler traffic junction macro models having single input-output lanes for each arm. This highlights the advantages of the EM algorithm when applied to traffic flow macro models as explored in this work.

Future work could address improvement of computational efficiency. In the proposed algorithm, uniform windows were applied, which look back in time and move on a time grid dictated by uniform time lags. Hence, the estimation algorithm is not strictly a real-time methodology because it requires measurement data to be stored for those time points falling inside the window frames. Improvements could be developed by using the measurement data only once and without storage (Elliott and Krishnamurthy, 1999), leading to a full real-time algorithm. To address the real time integration of control, communications and computational technologies, future work could investigate the reduction of strategic sensors per arm by exploiting the increasing availability of vehicle information from car-to-car and car-to-infrastructure communication.

REFERENCES

Astrom, K. J. and Wittenmark, B. (1995). *Adaptive Control*. Addison-Wesley, New York.

Beckmann, M., McGuire, C. B., and Winsten, C. B. (1955). *Studies in the Economics of Transportation*. The Cowles, Commission, Yale University Press.

Bishop, C. M. (2009). *Pattern Recognition and Machine Learning*. Springer, New York.

Chen, S. (2006). The Application of the Expectation-Maximization Algorithm to the Identification of Biological Models. Master's thesis, Faculty of the Virginia Polytechnic Institute and State University.

Dang, X. H., Lee, V. C. S., Ng, W. K., Ciptadi, A., and Ong, K. L. (2009). An EM-Based Algorithm for Clustering Data Streams in Sliding Windows. In *Proceedings of the 14th International Conference on Database Systems for Advanced Applications, Australia, April 21-23*.

Dempster, A., Laird, N. M., and Rubin, D. B. (1977). ML from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B 39, 1-38*.

Dunn Engineering Associates and Siemens Intelligent Transporation Systems (2005). Traffic Control Systems Handbook. Technical report, Federal Highway Administration.

Elliott, R. J. and Krishnamurthy, V. (1999). New Finite-Dimensional Filters for Parameter Estimation of Discrete-Time Linear Gaussian Models. *IEEE Transactions on Automatic Control, 44(5):938-951*.

Gazis, D. C. (2002). *Traffic Theory*. Kluwer Academic Publishers, New York.

Gipps, P. G. (1981). A behavioural car following model for computer simulation. *Transp. Res. B, 15, 403-414*.

Homolova, J. (2005). Traffic Flow Control. In *Proceedings of the 15th International Conference on Process Control, Slovak University of Technology, Bratislava, June 7-10*.

Huber, M. (2015). *Nonlinear Gaussian Filtering: Theory, Algorithms, and Applications*. KIT Scientific Publishing, Karlsruhe Institute of Technology.

Kratochvilova, J. and Nagy, I. (2004). Local traffic control of a microregion. *Ministry of Transportation, Czech Republic, National Programme of Research 2004-2009, Projectnum. 1F43A/003/120*.

Lighthill, M. J. and Whitham, G. B. (1955). On kinematic waves. II. A theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society of London, 229, 317-345*.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice-Hall, Inc, Sweden.

Mihaylova, L., Boel, R., and Hegyi, A. (2007). Freeway traffic estimation within particle filtering framework. *Automatica, 43, 290-300*.

Olszewski, P. S. (1994). Modeling probability distribution of delay at signalized intersections. *Journal of advanced transportation, 28:3, 253-274*.

Pecherkova, P., Dunik, J., and Flidr, M. (2008). *Robotics Automation and Control*, chapter 17, Modelling and Simultaneous Estimation of State and Parameters of Traffic System. InTech.

Richards, P. I. (1956). Shock waves on the highway. *Operations Research 4, 42-51*.

Sarkka, S. (2013). *Bayesian Filtering and Smoothing*. Cambridge University Press, Cambridge.

Soderstrom, T. and Stoica, P. (1989). *System Identification*. Prentice-Hall, USA.

Viti, F. and van Zuylen, H. J. (2004). Modeling Queues At Signalized Intersections. In *Proceedings of the 83rd Annual Meeting of the Transportation Research Board, Washington D.C., January 11-15*.

Wang, Y., Papageorgiou, M., and Messmer, A. (2008). Real-time freeway traffic state estimation based on extended Kalman filter: Adaptive capabilities and real data testing. *Transportation Research Part A, 42, 1340-1358*.

Webster, F. V. (1957). Traffic Signal Settings. Technical report, Road Research Technical Paper, No. 39, Road Research Laboratory, London.

Zammit, L. C., Fabri, S. G., and Scerri, K. (2016). Joint state and parameter estimation for a macro traffic junction model. In *Proceedings of the 24th Mediterranean Conference on Control and Automation, Greece, June 21-24*.