Proceedings Paper:

Hawes, M., Amer, H. and Mihaylova, L.S. orcid.org/0000-0001-5856-2223 (2017) Traffic State Estimation via a Particle Filter Over a Reduced Measurement Space. In: 2017 20th International Conference on Information Fusion (Fusion). 20th International Conference on Information Fusion, 10-13 Jul 2017, Xi'an, China. IEEE . ISBN 978-0-9964-5270-0

https://doi.org/10.23919/ICIF.2017.8009804

Traffic State Estimation via a Particle Filter Over a Reduced Measurement Space

Matthew Hawes, Hayder M Amer, Lyudmila Mihaylova

Department of Automatic Control and Systems Engineering, University of Sheffield, S1 3JD, UK

{m.hawes,hmamer1,l.s.mihaylova}@sheffield.ac.uk

***Abstract*—Traffic control and vehicle route planning require accurate estimates of the traffic state in order to be successfully implemented. This estimation problem can be solved by using particle filters in conjunction with macroscopic traffic models such as the stochastic compositional model. The accuracy of the estimates can be decreased for road segments where there are no measurements available. However, the inclusion of measurements for all segment boundaries carries a computational cost associated with the evaluation of the likelihood function required by the particle filter. To solve this problem, this paper proposes using the column based matrix decomposition method to select the most significant locations in the road network. This results in the particle filter being applied over a reduced measurement space, allowing a trade-off between computational efficiency and estimation accuracy to be achieved. A performance evaluation based on a simulated stretch of road is provided to validate the proposed method. It shows that by selecting half the original number of measurements, the computational time is reduced by approximately 9% without significantly decreasing the estimation accuracy. A more significant improvement in terms of savings in computational complexity can be expected when considering larger urban road networks.**

I. Introduction

There is an increasing number of vehicles on road networks causing increased congestion problems. In order to alleviate these problems various traffic control, [1], [2], and route planning, [3]–[5], methods have been proposed. However, for these methods to work they require an accurate estimate of the current traffic state.

Traffic state estimation involves modeling the road network, which is a complex problem with many interacting components and random perturbations [6]–[8]. For example, consider drivers in a traffic jam. As drivers approaching an incident observe the congestion forming in front of them they begin to slow down. The drivers following them see this change in speed and react in turn, resulting in a reduction in speed moving further up the road, away from the original incident.

There are three broad levels of models that can be used for this task: microscopic models that deal with individual vehicles [9], macroscopic models that consider aggregated measurements (flow and speed) [1], [8], [10]–[14] and mesoscopic models, which can be considered as a combination of the two, [15]. The ideas behind microscopic models can also be extended to consider platoons or groups of vehicles in an attempt to improve their computational efficiency [16]. However, due to their lower computational requirements macroscopic traffic models are often used in real time applications [7].

A common macroscopic model is the cell transmission model (CTM) [17], [18]. In the CTM a length of road is split into a sequence of links, each of which is further broken into smaller road segments called cells. The interactions between neighboring cells is then modeled by sending and receiving functions, which along with a maximum number of vehicles allowed in each cell controls the movement of vehicles between cells.

In [6] a flexible stochastic compositional model (SCM) is presented for online modeling of traffic flows. Here, a dynamic equation is used to describe the evolution in time of traffic speeds in each cell. It is flexible in terms of cell and time update sizes, with both being able to vary in time if required (as along as no vehicles completely skip a cell during a time step). The random nature of traffic state evolution can also be explicitly accounted for via probability distributions that govern the sending and receiving functions as well as noise terms.

When combined with such models Kalman filters (KFs) can be used to recursively estimate the traffic states [19]–[22]. Alternatively particle filters (PFs), [23], [24], have also been successfully applied to traffic estimation problems [7], [9] and shown to be powerful and scalable. These methods use observations up to the current point in time, along with system dynamics, to obtain the conditional distribution of the traffic state.

Although they can handle there not being measurements at every road segment boundary, it has been observed that the estimates they provide are more accurate at the boundaries which do have measurements present [7]. The temptation then is to ensure there are measurements available at each of the boundaries in order to improve the overall estimation accuracy. However, this means that more measurements are used evaluating the likelihood terms, thus increasing computational complexity. This leaves the question of how many of the measurements should be used and what is the best way to select them.

One method of representing a road network in a compressed form would be to use principal component analysis (PCA)

[25]–[27]. This compressed form is given as basis vectors and latent variables. However, it can be difficult to assign a physical meaning to the latent variables, making them hard to interpret. Additionally, all the information from the traffic sensors being considered have to be collected at each point in time during real world application.

The authors of [28] use column based matrix decomposition, [29]–[31], in order to give the overall road network in terms of a smaller subnetwork. This involves the singular value decomposition (SVD) of the matrix containing the measurements for the entire traffic network. The locations in the network with the highest variations in measurements are then kept with a higher probability and used to approximate the network as a whole. This scheme will not outperform PCA in terms of compression. However, the matrices involved can easily be interpreted and lend themselves to use within a PF as a result. Additionally, only measurements for the retained locations have to be made in real world application rather than recording all of the information from each of the traffic sensors.

This paper proposes using the column based matrix decomposition to select which segment boundary measurements should be used in the evaluation of the likelihood function in a PF. As a result, there is a reduced measurement space. The estimate of the overall traffic state is then provided by the PF in conjunction with the SCM. A performance evaluation on a simulated stretch of road is provided to validate the proposed traffic state estimation method. It is reasonable to expect greater performance improvements to be found in higher dimension problems (urban environments), where there are more measurements available to begin with [32].

The remainder of this paper is structured as follows: Section II gives details of the traffic model used. This includes the details of the SCM (II-A) and the measurements model (II-B). Then in Section III the proposed PF for a reduced measurement space is detailed, including the measurement selection (III-A) and the overall PF framework (III-B). A performance evaluation is provided in Section IV and finally concluding remarks are given in Section V.

II. Traffic Flow Model

A. Traffic Model

This paper considers the SCM [6], where the road is split into segments or cells as shown in Figure 1. Here L_i is the length of road segment i and segment i consists of l_i lanes. The task is then to estimate the traffic states, given by $\mathbf{x}_k = [\mathbf{x}_{1,k}^T, \mathbf{x}_{2,k}^T, ..., \mathbf{x}_{n,k}^T]^T$, $\mathbf{x}_{i,k} = [N_{i,k}, v_{i,k}]^T$, where $N_{i,k}$ and $v_{i,k}$ are the number of vehicles and their average speed, respectively, at K times $t_1, t_2, ..., t_k, ...t_K$. Note, n gives the number of road segments, segment $n+1$ is the fictitious last road segment and the average vehicle length is assumed to be A_l.

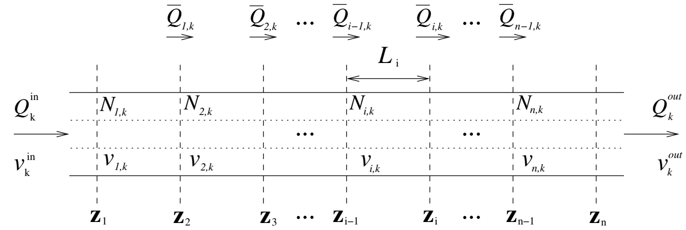

Fig. 1: Road segments and measurement points.

The following equations describe the evolution of the traffic states:

$$\mathbf{x}_{1,k+1} = f_1(Q_k^{in}, v_k^{in}, \mathbf{x}_{1,k}, \mathbf{x}_{2,k}, \boldsymbol{\eta}_{1,k}), \quad (1)$$

$$\mathbf{x}_{i,k+1} = f_i(\mathbf{x}_{i-1,k}, \mathbf{x}_{i,k}, \mathbf{x}_{i+1,k}, \boldsymbol{\eta}_{i,k}), \quad (2)$$

$$\mathbf{x}_{n,k+1} = f_n(\mathbf{x}_{n-1,k}, \mathbf{x}_{n,k}, Q_k^{out}, v_k^{out}, \boldsymbol{\eta}_{n,k}), \quad (3)$$

where f_i is specified by the traffic model and $\boldsymbol{\eta}_k$ allows for random fluctuations and modeling errors. Here, Q_k^{in} and Q_k^{out}, are the number of vehicles entering the first segment and leaving the last segment within the time interval $\Delta t_k = t_{k+1} - t_k$ with average speeds v_k^{in} and v_k^{out}, respectively. These values are boundary conditions required by the filter, rather than states to be estimated.

The traffic behavior is modeled with forward and backward propagation of traffic perturbations. This is achieved by finding the sending ($S_{i,k}$) and receiving ($R_{i,k}$) functions, which give the number of vehicles that can leave and enter each road segment, respectively. The model also finds/makes use of the anticipated traffic density ($\rho_{i,k+1}^{antic}$) as a result of mixing densities from two neighboring cells, a threshold density (ρ_{th}) and the intermediate velocity ($v_{i,k+1}^{interm}$), which can be viewed as a kind of mixing velocities from neighboring cells. Further details on the SCM for traffic flow can be found in [6]. Note, this model could be extended to model urban road networks by considering the turning fractions of vehicles at junctions or by using origin-destination information for the vehicles if this available [18].

B. Measurement model

In this work the measurements of interest are the flow of vehicles past segment boundaries ($\bar{Q}_{j,s}$) and their associated average speeds ($\bar{v}_{j,s}$). Traditionally, such measurements can be obtained using induction loops under the road surface, and more recently from radar, video cameras or global positioning system (GPS) on probe vehicles. The measurements available (i.e. measurements for all of the n boundaries) at time t_s are given by $\mathbf{z}_s = [\mathbf{z}_{1,s}^T, \mathbf{z}_{2,s}^T, ..., \mathbf{z}_{n,s}^T]^T$, where $\mathbf{z}_{j,s} = [\bar{Q}_{j,s}, \bar{v}_{j,s}]^T$.

The measurement equation is given by:

$$\mathbf{z}_s = h(\mathbf{x}_s, \boldsymbol{\xi}_s), \quad (4)$$

where $h(.)$ is determined by the measurement model used. If a Gaussian measurement noise is assumed this gives:

$$\mathbf{z}_{j,s} = \begin{pmatrix} \bar{Q}_{j,s} \\ \bar{v}_{j,s} \end{pmatrix} + \boldsymbol{\xi}_s, \quad (5)$$

where $\boldsymbol{\xi}_s = [\xi_{\bar{Q}_{j,s}}, \xi_{\bar{v}_{j,s}}]^T$. Therefore, from a known distribution of the initial state vector the estimation problem discussed in Section II-A becomes a recursive Bayesian estimation problem. This can be solved using a PF as detailed in Section III-B.

III. Particle Filtering Framework for Traffic State Estimation with Reduced Measurement Space

A. Measurement Selection

This section details how the column based matrix decomposition can be used to select the measurements (flows across segment boundaries and average vehicle speeds) to give a reduced measurement space, to which a PF can be applied for traffic state estimation. First assume that $\mathbf{Z}_a \in \mathcal{R}^{2K_a \times n}$ is the matrix containing all of the available the measurements for all points in the road network for a given period of time. Also, K_a gives the number of points in time considered in the matrix $\mathbf{Z}_a$ being decomposed. The dimension of interest is twice this length as the measurements being considered contain both flow measurements and the associated average speeds of the vehicles.

This measurement matrix can then be approximated by $\hat{\mathbf{Z}}$, which is given by [28]:

$$\hat{\mathbf{Z}} = \tilde{\mathbf{Z}}\boldsymbol{\Phi}. \tag{6}$$

Here, $\tilde{\mathbf{Z}} \in \mathcal{R}^{2K_a \times m}$ gives the measurements at all points in time for a subset of the road segment boundaries, $\boldsymbol{\Phi} \in \mathcal{R}^{m \times n}$ express the columns in $\hat{\mathbf{Z}}$ in terms of the basis given in $\tilde{\mathbf{Z}}$. Note, as $\tilde{\mathbf{Z}}$ only contains a subset of all of the available measurements this means that $m < n$.

To find the m locations from which the measurements are used to construct $\tilde{\mathbf{Z}}$ the SVD of $\mathbf{Z}_a$ is found. The right singular vector is then used to assign a probability (P_{z_i}) of each location/column being selected. This probability is given by

$$P_{z_i} = \frac{1}{r}\sum_{j}^{r} \check{v}_{i,j}, \quad \text{for } i = 1, 2, ..., n, \tag{7}$$

where r is the rank of $\mathbf{Z}_a$ and $\check{v}_{i,j}$ the i^{th} entry in the j^{th} right singular vector. Once these probabilities have been found the m locations with the highest proabilities can then be kept and used to approximate the road network as a whole.

In [28] the authors discuss learning the matrix $\boldsymbol{\Phi}$ from the available data. Then when new measurements (now only for the m retained locations) become available the new measurements for the road network as a whole can be given by

$$\hat{\mathbf{Z}}_{new} = \tilde{\mathbf{Z}}_{new}\boldsymbol{\Phi}. \tag{8}$$

Here, $\tilde{\mathbf{Z}}_{new}$ is the newly available measurements and $\hat{\mathbf{Z}}_{new}$ the resulting approximation of what the road network is doing as a whole.

Instead, this work constructs a new measurement vector, $\check{\mathbf{z}}_k = [\check{\mathbf{z}}_{n_1,k}^T, \check{\mathbf{z}}_{n_2,k}^T, ..., \check{\mathbf{z}}_{n_m,k}^T]^T$, when the measurements for the m selected locations become available at time t_k. This process is summarised in Figure 2, where the selected locations are given by $n_1, n_2, ..., n_m$. This measurement vector is then used as the measurement vector in the particle filter as detailed below. Estimates for the state at each location are then provided by the PF rather than the approximation scheme in [28].

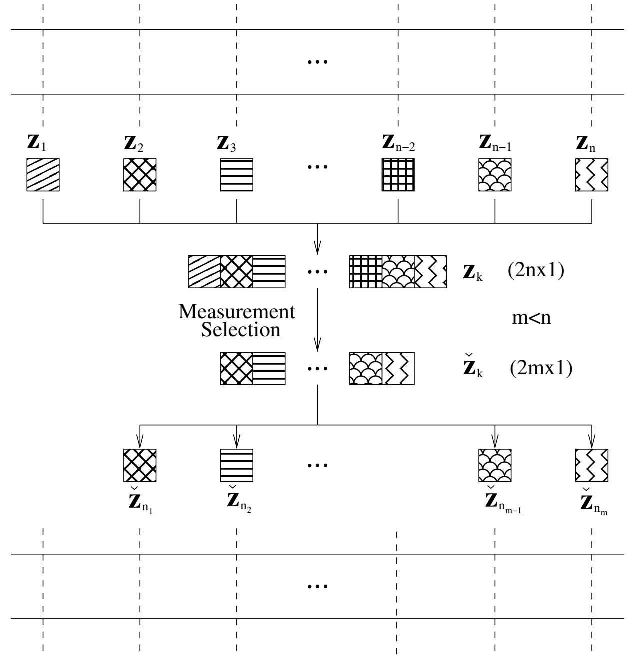

Fig. 2: Framework for construction of reduced measurement space.

B. Particle Filtering Framework

In PFs, the aim is to find the posterior probability density function (PDF) of the state at time t_k given a set of measurements up to the same point in time. This involves evaluating $p(\mathbf{x}_k|\check{\mathbf{Z}}^k)$, where $\check{\mathbf{Z}}^k = [\check{\mathbf{z}}_1^T, ..., \check{\mathbf{z}}_k^T]^T$ and $\check{\mathbf{z}}_k$ is constructed as detailed above.

From Bayes' rule

$$p(\mathbf{x}_k|\check{\mathbf{Z}}^k) = \frac{p(\check{\mathbf{z}}_k|\mathbf{x}_k)p(\mathbf{x}_k|\check{\mathbf{Z}}^{k-1})}{p(\check{\mathbf{z}}_k|\check{\mathbf{Z}}^{k-1})}, \tag{9}$$

where

$$p(\mathbf{x}_k|\check{\mathbf{Z}}^{k-1}) = \int_{\mathbb{R}^{n_x}} p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\check{\mathbf{Z}}^{k-1})d\mathbf{x}_{k-1} \tag{10}$$

and $p(\check{\mathbf{z}}_k|\check{\mathbf{Z}}^{k-1})$ is a normalising constant. This means $p(\mathbf{x}_k|\check{\mathbf{Z}}^k)$ can be updated using the following proportionality relationship:

$$p(\mathbf{x}_k|\check{\mathbf{Z}}^k) \propto p(\check{\mathbf{z}}_k|\mathbf{x}_k)p(\mathbf{x}_k|\check{\mathbf{Z}}^{k-1}). \tag{11}$$

However, this process is computationally expensive, meaning it is necessary to use methods such as the PF that give approximate solutions [23], [24]. Algorithm 1 gives the PF (with M_{pf} particles) for traffic state estimation with reduced measurement space that is considered in this work. The

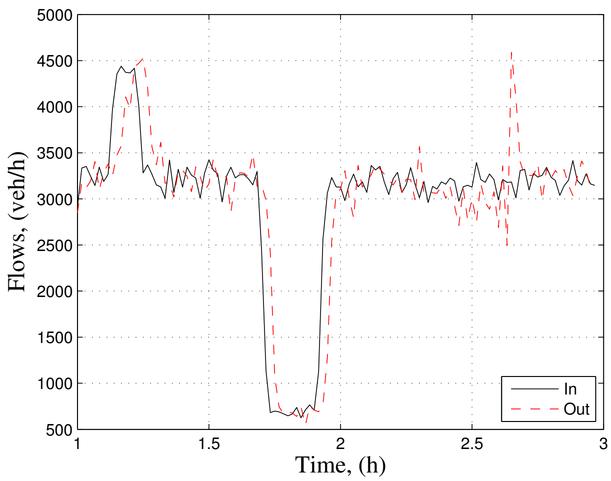

Fig. 3: Flow of vehicles at inflow and outflow road boundaries.

interested reader is referred to [7] for further details on the use of PFs for traffic state estimation.

The novelty of this work is in the inclusion of the measurements selection step (Step 1 in Algorithm 1), the results of which form a reduced measurement space (Step 4 in Algorithm 1). Using this step reduces the number of measurements required to evaluate the likelihood function in the PF (Step 5 in Algorithm 1), which in turn reduces the computational complexity. The measurement selection step, which gives the reduced measurements space is detailed above in Section III-A.

Note, $t_k \equiv t_s$ is required to account for the fact that there is not necessarily measurements available at every time step within the particle filter. For example it is possible to consider a time step of ten seconds (model applied every ten seconds) but only have measurements available every minute.

IV. Performance Evaluation

In this section a performance evaluation based on a 4km stretch of simulated road with 3 lanes is provided for a 3 hour period of time. Congestion was introduced by varying the inflow and outflow rates for the simulated road, as shown in Figure 3. The corresponding average speeds are shown in Figure 4. Multiple sets of measurements are generated, one is then used for the measurement selection step as detailed in Section III-A and the remaining for the $MC = 100$ runs. The road segments were initialised to have 14 vehicle present, with an average speed of 100km/h. The interested reader is referred to [7] for further details about how the simulated data used for testing was generated.

A comparison is drawn between PFs with a varying number of measurements being utilised. Firstly, with all available measurements being used. Then with $m = 4$ measurements being selected by the proposed method. Note, 4 measurements are used as an example here to illustrate that it is possible to make a computational saving. However, as a general rule it

Algorithm 1 Particle Filter for Traffic State Estimation with Reduced Measurement Space (adapted from [7])

1: Measurement Selection: Carry out the decomposition to get an approximation of the road network

$$\hat{\mathbf{Z}} = \tilde{\mathbf{Z}}\mathbf{\Phi}$$

to select the m most significant locations.

2: Initialization: $k = 0$
   For $l = 1, \ldots, M_{pf}$
      Generate samples $\{\mathbf{x}_0^{(l)}\}$ from the initial distribution $p(\mathbf{x}_0)$ and initial weights $w_0^{(l)} = 1/M_{pf}$.
   End For

3: Prediction step:
   For $l = 1, \ldots, M_{pf}$,
      sample $\mathbf{x}_k^{(l)} \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(l)})$ according to Section II-A for the road segments of interest.
   End For

4: Construct the vector for the reduced measurement space, only for $t_k \equiv t_s$, $\check{\mathbf{z}}_k = [\check{\mathbf{z}}_{n_1,k}^T, \check{\mathbf{z}}_{n_2,k}^T, ..., \check{\mathbf{z}}_{n_m,k}^T]^T$.

5: Measurement processing step, only for $t_k \equiv t_s$: Compute the weights
   For $l = 1, \ldots, M_{pf}$
      $w_s^{(l)} = w_{s-1}^{(l)} p(\check{\mathbf{z}}_s|\mathbf{x}_s^{(l)})$,
   End For
   where the likelihood $p(\check{\mathbf{z}}_s|\mathbf{x}_s^{(l)})$ is determined by the measurement model in Section II-B.
   For $l = 1, \ldots, M_{pf}$
      Normalize the weights: $\hat{w}_s^{(l)} = w_s^{(l)} / \sum_{l=1}^{M_{pf}} w_s^{(l)}$.
   End For

6: Output: $\hat{\mathbf{x}}_s = \sum_{l=1}^{M_{pf}} \hat{w}_s^{(l)} \mathbf{x}_s^{(l)}$,

7: Selection step (resampling), only for $t_k \equiv t_s$:
   Multiply/ Suppress samples $\mathbf{x}_s^{(l)}$ with high/ low importance weights $\hat{w}_s^{(l)}$, in order to obtain M random samples approximately distributed according to $p(\mathbf{x}_s^{(l)}|\check{\mathbf{Z}}^s)$, e.g. by residual resampling.
   For $l = 1, \ldots, M_{pf}$,
      $w_s^{(l)} = \hat{w}_s^{(l)} = 1/M_{pf}$,
   End For

8: $k \leftarrow k + 1$ and return to step (1).

would be reasonable to expect a larger computation saving when even less measurements are used. This would be at the cost of reduced accuracy in terms of the estimations made. A greater sensetivity analysis for the proposed method is currently being undertaken for a larger urban traffic network.

Estimation accuracy is illustrated by using the root mean square error ($RMSE$) as given by

$$RMSE_{\epsilon_{i,j,k}} = \sqrt{\frac{\sum_{j=1}^{MC}\sum_{i=1}^{n}(\epsilon_{i,j,k} - \hat{\epsilon}_{i,j,k})^T(\epsilon_{i,j,k} - \hat{\epsilon}_{i,j,k})}{MC \times n}}, \quad (12)$$

where $\epsilon_{i,j,k}$ is the actual value of interest and and $\hat{\epsilon}_{i,j,k}$ the

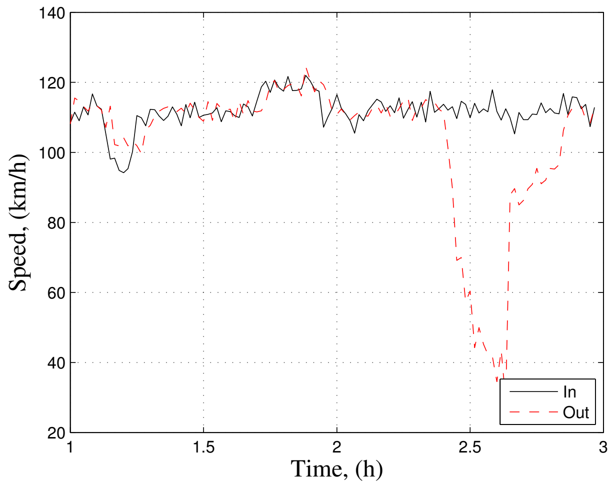

Fig. 4: Average vehicle speeds at inflow and outflow road boundaries.

corresponding estimated value of interest. In this work the parameters for which an $RMSE$ value is found for are: the density of the vehicles inside the cells ($\rho_{i,k} = N_{i,k}/(L_i \times l_i)$), the flow of the vehicles past segment boundaries ($\bar{Q}_{i,k}$) and the vehicle speeds ($v_{i,k}$).

Computational efficiency will be judged by considering the computation time. Note, the time for implementing the column based decomposition for measurement selection is included in the time given for the example where the measurements are being selected by the proposed method. All results were obtained in Matlab on a computer with an Intel Xeon CPU E3-1271 (3.60GHz) and 16GB of RAM.

In all cases the values $M_{pf} = 100$, $\sigma^2_{\xi_{Q_{j,s}}} = 3(\text{veh})^2$ and $\sigma^2_{\xi_{v_{j,s}}} = 2(\text{km/h})^2$ were used. Finally the remaining parameter values required by the particle filter and the traffic model are as follows: $v_{free} = 120km/h$, $v_{min} = 7.4km/h$, $\rho_{crit} = 20.89veh/km/lane$, $\rho_{jam} = 180veh/km$, $A_l = 0.01km$. The PF with varying numbers of measurements will be compared for all road locations and an example individual location (sixth) in what follows.

A. Overall Comparison

Here the results for the PF with measurements available at all segment boundaries and at four selected measurement boundaries are summarised in Table I. The $RMSE$ values and computation times are given as mean values over the entirety of the time period of interest.

Firstly, it can be seen that decreasing the number of measurements used to evaluate the likelihood function has decreased the computation time required by the PF (a reduction of 9.06%). Recent work considering the big/tall data problem with Markov chain Monte Carlo based methods has suggested that the improvement in terms of computational savings is greater for higher dimension problems, where there are originally more measurements available to begin with

TABLE I: Performance summary for PF with 8 measurements and 4 measurements.

Example	8 Measurements	4 Measurements
$\overline{RMSE}_\rho$ (veh/km)	5.37	5.59
$\overline{RMSE}_{\bar{Q}}$ (veh/h)	798.46	801.53
$\overline{RMSE}_v$ (km/h)	12.10	13.05
Computation Time (minutes)	23.83	21.67

[32]. As a result, it would be reasonable to expect a more significant reduction in terms of computational complexity when considering the problem of traffic state estimation for urban environments.

This can be further illustrated if by considering the fact that the difference in computational complexity can be given as:

$$\begin{aligned} \mathcal{O}(nM_{pf} - mM_{Pf}) &= \mathcal{O}(nM_{pf} - cnM_{pf}) \\ &= \mathcal{O}(nM_{pf}(1-c)), \end{aligned} \quad (13)$$

where n is the total possible number of measurements, m the number of measurements selected by the decomposition method and c the ratio of measurements kept. Therefore, if the number of particles and ratio of measurements used by the PF is kept constant, increasing the number of potential measurements that can be used gives a larger potential saving in computational complexity. Increasing the number of particles can therefore also have a similar effect.

From Table I it can also be seen that the $RMSE$ values have increased for each value being estimated when less measurements are used. This suggests that the resulting estimates are less accurate then when all of the measurements are used. However, despite the increases in $RMSE$ values an acceptable estimation accuracy is still achieved.

To illustrate the estimation accuracy is still acceptable consider Figures 5-6. These figures show example performances for a single run of the PF using only the four measurements. Figures 5-6 show the spatio-temporal evolution of the traffic. The colour bar represents the number of vehicles crossing the segment boundaries (or their associated speeds). Here, it can be seen that there the number of vehicles crossing segment boundaries and the associated speeds have been estimated with a good accuracy. Note, the reduction in flow shown in Figure 5 corresponds to the reduction in inflow and outflow between the times of 1.70h and 1.82h shown in Figure 3. Similarly the reduction in speeds shown in Figure 6 is explained by the large reduction in outflow speed indicated in Figure 4. In addition the flow density diagram in Figure 7 shows the usual characteristics associated with traffic flow, further indicating that appropriate estimates have been achieved. Note, the colours in Figure 7 are used to distinguish between different measurement locations.

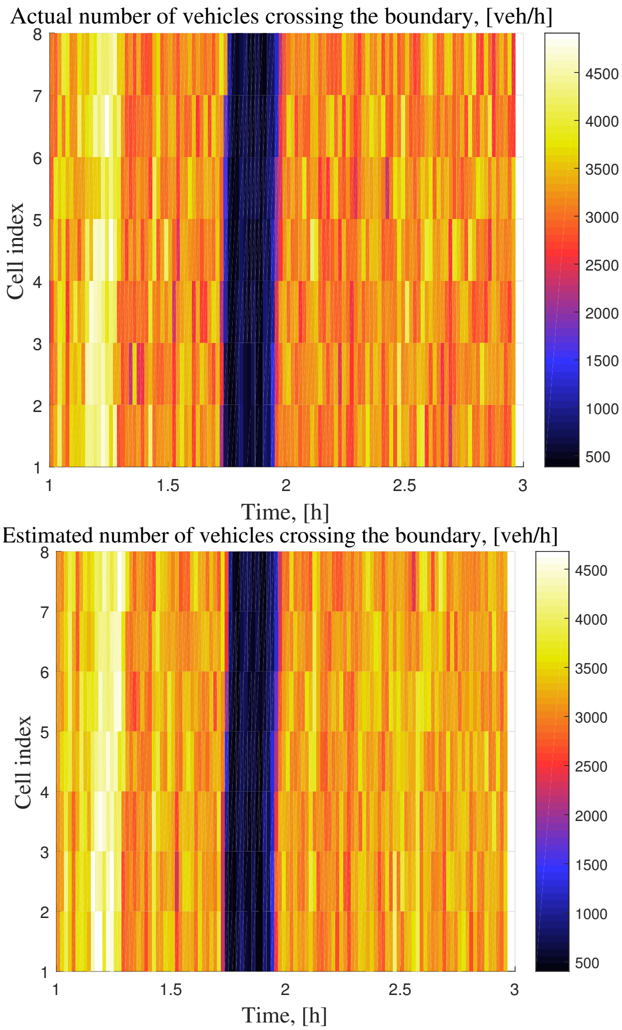

Fig. 5: Actual and estimated numbers of vehicles crossing cell boundaries.

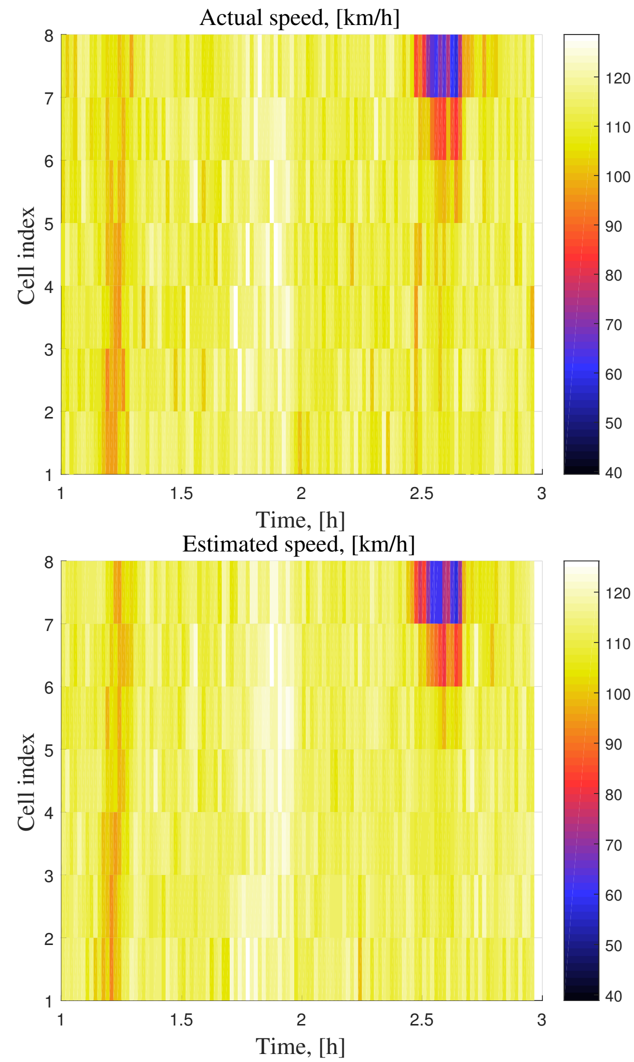

Fig. 6: Actual and estimated numbers of vehicles crossing cell boundaries.

B. Comparison for the sixth Segment Boundary

Here, the sixth road segment will be considered to illustrate the performance in terms of estimation accuracy for a single location on the road. The $RMSE$ values are illustrated in Figures 8-10 and summarised in Table II. Note, computation times are not shown here as they are illustrated in the overall comparison above. It can be seen that there are acceptable estimation accuracies in both cases. Although the $RMSE$ values are higher when only four measurements are used, the increase has not been significant enough to give inappropriate estimates.

V. Conclusions

Traffic state estimation is an important first step in solving problems such as route planning for congestion avoidance and traffic control measures. Particle filters have been shown to be a powerful method of solving this estimation problem. This paper proposed using the column based matrix decomposition method to select the measurements used to construct a reduced

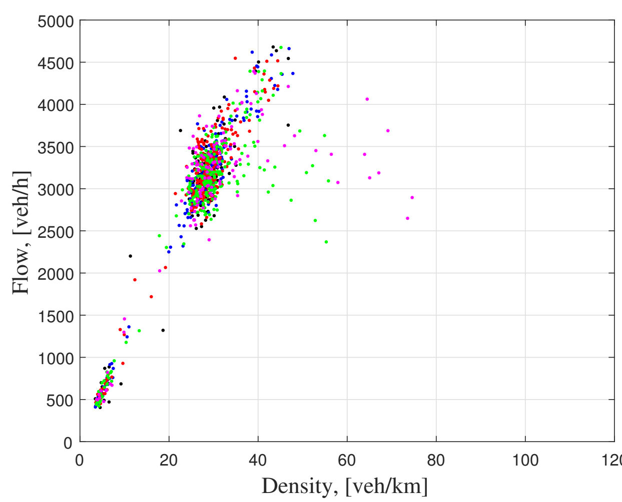

Fig. 7: Example-flow density diagram.

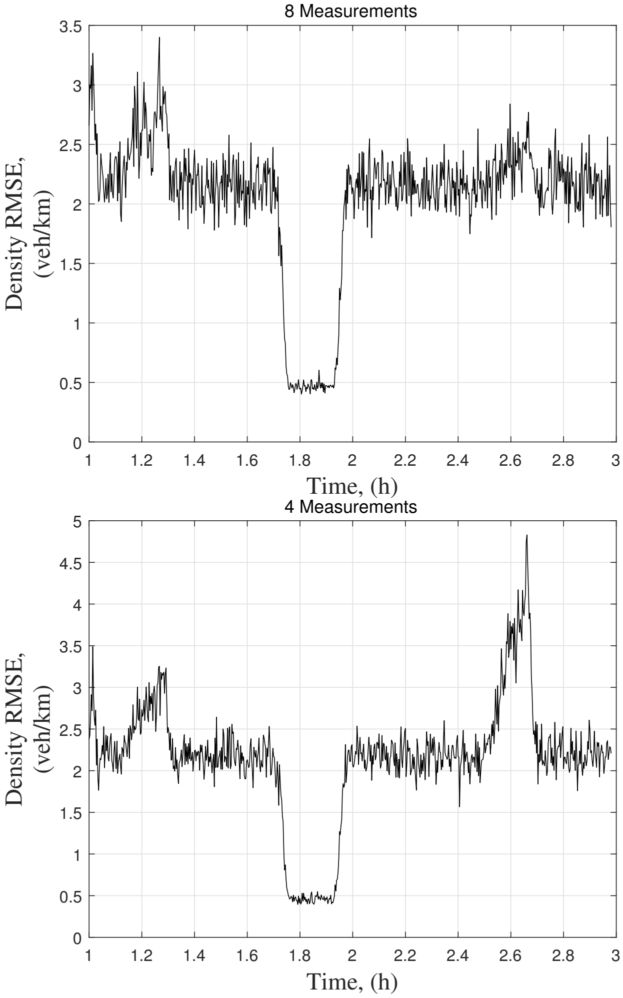

Fig. 8: $RMSE$ for the density estimates at thesixth road segment.

TABLE II: Performance summary at the sixth road segment for a PF with 8 measurements and 4 measurements.

Example	8 Measurements	4 Measurements
$\overline{RMSE}_\rho$ (veh/km)	2.04	2.14
$\overline{RMSE}_{\bar{Q}}$ (veh/h)	297.01	299.37
$\overline{RMSE}_v$ (km/h)	4.32	5.18

measurement space for use within the particle filter for traffic state estimation. A performance evaluation is provided for a simulated stretch of road and shows that an 9.06% improvement in terms of computational time is possible when selecting half the original number of measurements to use. This has not come at the cost of a significant decrease in estimation accuracy. It would also be reasonable to expect a more significant improvement in terms of reduction in computational complexity when considering large urban road networks.

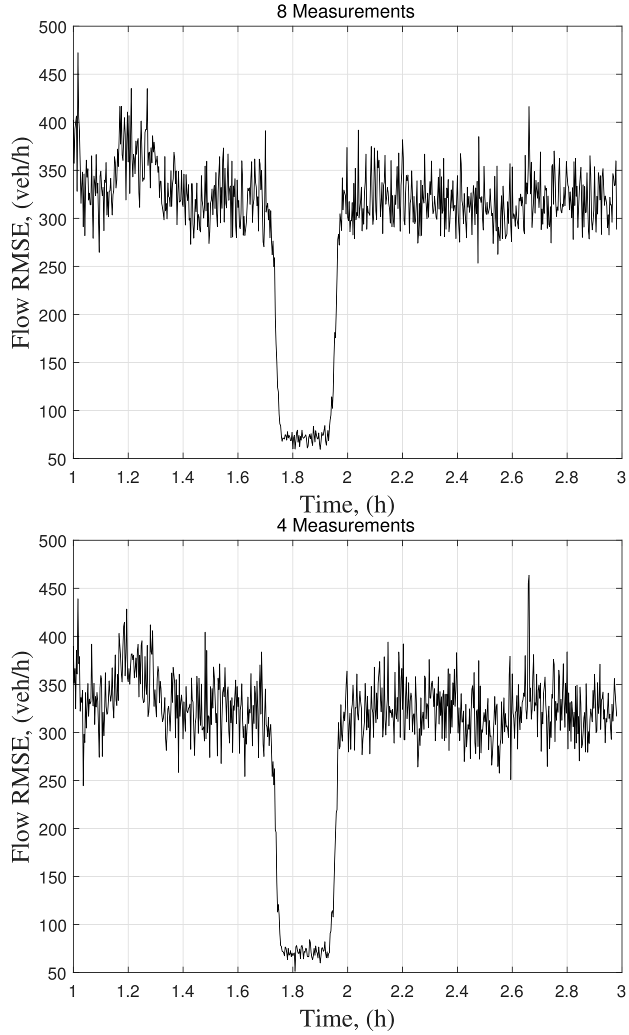

Fig. 9: $RMSE$ for the flow estimates at the sixth road segment.

Acknowledgments

The authors appreciate the support of the SETA project funded from the European Union's Horizon 2020 research and innovation program under grant agreement no. 688082 and the Iraqi Government, especially The Ministry of Higher Education and The Committee for Education Development in Iraq.

References

[1] M. van den Berg, A. Hegyi, B. D. Schutter, and J. Hellendoorn, "A macroscopic traffic flow model for integrated control of freeway and urban traffic networks," in *Proceedings of the IEEE Conference on Decision and Control*, 2003, pp. 2774–2779.

[2] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2043–2067, 2003.

[3] T. N. Nha, V., S. Djahe, and J. Murphy, "A comparitive study of vehicles' routing algorithms for route planning in smart cities," in *Proceedings of the First International Workshop on Vehicular Traffic Management for Smart Cities*, 2012, pp. 1–6.

[4] H. Amer, N. Salman, M. Hawes, M. Chaqfeh, L. Mihaylova, and M. Mayfield, "An improved simulated annealing technique for enhanced mobility in smart cities," *Sensors*, vol. 16, no. 7, p. 1013, 2016.

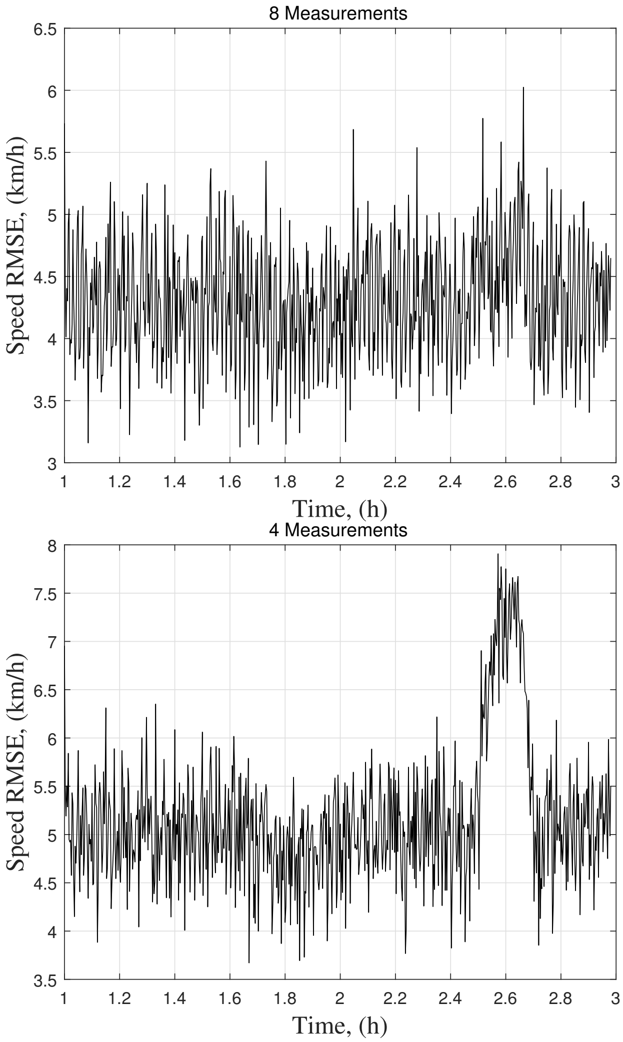

Fig. 10: $RMSE$ for the speed estimates at the sixth road segment.

[5] A. Faro and D. Giordano, "Algorithms to find shortest and alternative paths in free flow and congested traffic regimes," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 7, pp. 441–444, 2006.

[6] R. Boel and L. Mihaylova, "A compositional stochastic model for real time freeway traffic simulation," *Transportation Research Part B: Methodological*, vol. 40, no. 4, pp. 319 – 334, 2006.

[7] L. Mihaylova, R. Boel, and A. Hegyi, "Freeway traffic estimation within particle filtering framework," *Automatica*, vol. 43, no. 2, pp. 290 – 300, 2007.

[8] M. Treiber and A. Kesling, *Traffic Flow Dynamics: Data, Models and Simulation*. Berlin-Heidelberg: Springer, 2013.

[9] S. Hoogendoorn, S. Ossen, M. Schreuder, and B. Gorte, "Unscented particle filter for delayed car-following models estimation," in *Proceedings of the IEEE Intelligent Transportation Systems Conference*, 2006, pp. 1598–1603.

[10] M. Papageorgiou, J.-M. Blosseville, and H. Hadj-Salem, "Macroscopic modelling of traffic flow on the Boulevard Peripherique in Paris," *Transportation Research Part B: Methodological*, vol. 23, no. 1, pp. 29 – 47, 1989.

[11] S. Hoogendoorn and P. Bovy, "Exploiting structure in wavelet-based Bayesian compressive sensing," *Journal of Systems Control Engineer-Proceedings of the Intnational Mechanical Engineers, Part I*, vol. 215(14), pp. 283–303, 2001.

[12] A. Kotsialos, M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham, "Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 282–292, 2002.

[13] S. Blandin, A. Couque, A. Bayen, and D. Work, "On sequential data assimilation for scalar macroscopic traffic flow models," *Physica D: Nonlinear Phenomena*, vol. 241, no. 17, pp. 1421 – 1440, 2012.

[14] S. Fan and D. B. Work, "A heterogeneous multiclass traffic flow model with creeping," *SIAM Journal on Applied Mathematics*, vol. 75, no. 2, pp. 813–835, 2015.

[15] M. D. Gangi, "Modeling evacuation of a transport system: Application of a multimodal mesoscopic dynamic traffic assignment model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1157–1166, 2011.

[16] N. E. Marinic, A. Sarlette, and R. K. Boel, "Distributed particle filter for urban traffic networks using a platoon-based model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1918–1929, 2013.

[17] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269 – 287, 1994.

[18] ——, "The cell transmission model, part ii: Network traffic," *Transportation Research Part B: Methodological*, vol. 29, no. 2, pp. 79 – 93, 1995.

[19] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended Kalman filter: a general approach," *Transportation Research Part B: Methodological*, vol. 39, no. 2, pp. 141 – 167, 2005.

[20] D. B. Work, O. P. Tossavainen, S. Blandin, A. M. Bayen, T. Iwuchukwu, and K. Tracton, "An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices," in *Proceedings of the IEEE Conference on Decision and Control*, 2008, pp. 5062–5068.

[21] L. L. Ojeda, A. Y. Kibangou, and C. C. de Wit, "Adaptive Kalman filtering for multi-step ahead traffic flow prediction," in *Proceedings of the American Control Conference*, 2013, pp. 4724–4729.

[22] P.-E. Mazar, O.-P. Tossavainen, and D. B. Work, "Computing travel times from filtered traffic states," *Discrete and Continuous Dynamical Systems - Series S*, vol. 7, no. 3, pp. 557–578, 2014.

[23] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings F - Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, 1993.

[24] A. Doucet, N. Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York, USA: Springer, 2001.

[25] Q. Li, H. Jianming, and Z. Yi, "A flow volumes data compression approach for traffic network based on principal component analysis," in *Proceedings of the IEEE Intelligent Transportation Systems Conference*, 2007, pp. 125–130.

[26] T. Djukic, J. Van Lint, and S. Hoogendoorn, "Application of principal component analysis to predict dynamic origin-destination matrices," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2283, pp. 81–89, 2012.

[27] M. T. Asif, S. Kannan, J. Dauwels, and P. Jaillet, "Data compression techniques for urban traffic data," in *Proceedings of the IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems*, 2013, pp. 44–49.

[28] N. Mitrovic, M. T. Asif, J. Dauwels, and P. Jaillet, "Low-dimensional models for compressed sensing and prediction of large-scale traffic data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2949–2954, 2015.

[29] P. Drineas, M. W. Mahoney, and S. Muthukrishnan, *Subspace Sampling and Relative-Error Matrix Approximation: Column-Based Methods*. Berlin, Heidelberg: Springer, 2006, pp. 316–326.

[30] C. Boutsidis, W. Mahoney, M., and P. Drineas, "Unsupervised feature selection for principle component analysis," in *Proceedings of the International Conference of Knowledge Discovery and Data Minning*, 2008, pp. 61–69.

[31] ——, "An improved approximation algorithm fort he column based subset selection problem," in *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorihtms*, 2009, pp. 968–977.

[32] A. D. Freitas, F. Septier, L. Mihaylova, and S. Godsill, "How can subsampling reduce complexity in sequential mcmc methods and deal with big data in target tracking?" in *Proceedings of the International Conference on Information Fusion*, 2015, pp. 134–141.