

# Adaptive Kalman Filtering for Multi-Step ahead Traffic Flow Prediction

Luis Ramon Leon Ojeda, Alain Y. Kibangou, Carlos Canudas de Wit

► **To cite this version:**

Luis Ramon Leon Ojeda, Alain Y. Kibangou, Carlos Canudas de Wit. Adaptive Kalman Filtering for Multi-Step ahead Traffic Flow Prediction. American Control Conference (ACC 2013), Jun 2013, Washington, DC, United States. hal-00842684

**HAL Id: hal-00842684**

**<https://hal.inria.fr/hal-00842684>**

Submitted on 9 Jul 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Adaptive Kalman Filtering for Multi-Step ahead Traffic Flow Prediction

Luis Leon Ojeda\*

Alain Y. Kibangou\*\*

Carlos Canudas de Wit\*\*\*

**Abstract**—Given the importance of continuous traffic flow forecasting in most of Intelligent Transportation Systems (ITS) applications, where every new traffic data become available in every few minutes or seconds, the main objective of this study is to perform a multi-step ahead traffic flow forecasting that can meet a trade-off between accuracy, low computational load, and limited memory capacity. To this aim, based on adaptive Kalman filtering theory, two forecasting approaches are proposed. We suggest solving a multi-step ahead prediction problem as a filtering one by considering pseudo-observations coming from the averaged historical flow or the output of other predictors in the literature. For taking into account the stochastic modeling of the process and the current measurements we resort to an adaptive scheme. The proposed forecasting methods are evaluated by using measurements of the Grenoble south ring.

## I. INTRODUCTION

Short-term traffic flow forecasting has drawn a lot of attention on the research field over the past decades due to the crucial and easily witnessed impact on daily life basis. Its importance can be seen in different traffic control applications, for instance [1], [2]. This subject has been amply investigated and several survey papers are available in the literature (see for example [3] and [4] and references therein). There exist several alternatives to categorize short-term traffic flow forecasting methodologies. One of them, as presented in [5], is from the standpoint of how information is used. Considering this, the literature can be divided into three groups: 1) approaches that make use of historical information exclusively: auto-regressive integrated moving average (ARIMA) family of models [6]–[8], non-parametric techniques, including non-parametric regression [6], [9]–[11], and neural networks [12]–[14]; 2) those using current information exclusively: strategies based on smoothing techniques [15]–[17], state-space models [18]–[20], and support vector machine [15]; and 3) those using both historical and current data: [5], [15], [21], [22]. In this framework, some works are worth noting. For instance the approach proposed in [5] uses, under a stochastic approach, both current and historical traffic data in an integrated way. On the other hand [21] uses a deterministic mixture of a constant predictor, defined as the difference between the historical and current flow levels at the starting time of prediction, and a nonparametric approach. This predictor succeeds to

capture current dynamics of the traffic flow. A more recent work, [22], proposes an on-line monitoring algorithm that, integrated with ARIMA models, can adapt itself to any potential new regimes. This result was met with greater accuracy than the basic ARIMA, but the main disadvantage is that an ARIMA model needs to be previously fitted.

Although an important part of the literature has been oriented to compare the performance of the different forecasting approaches, none has clearly pointed out the best solution. However, after its publication, the GML (Gaussian Maximum Likelihood) approach, proposed in [5], has been compared with other models, such as non-parametric regression, ARIMA, neural networks, online SVR, and smoothing techniques, concluding that GML-based predictor presented the best forecasting performance, under typical traffic conditions [15], [23], [24].

Apart from accuracy, another crucial matter in traffic flow forecasting is the development of approaches that forecast more than one step ahead in the future, as well as ones that can efficiently be implementable on-line. When it comes to multi-step ahead forecasting methodology, research shows that most authors prefer the use of approaches, such as neural networks [25]–[27], non-parametric regression [9], ARIMA model [28], and ATHENA model [29] (based on ARIMA model). Nevertheless, some other strategies can be used to achieve this goal, for instance, smoothing techniques [15], [21], GML-based predictor [5], and state-space based approaches [30].

Most of the forecasting results reviewed are based on offline computation. For online implementation, to the best of our knowledge, only few techniques have given rise to online implementation where low computational load and short time model calibration are required. Among these techniques, ARIMA models have been applied in different real scenarios. However, as [6] pointed out, the necessary model calibration for the seasonal ARIMA models was quite time-consuming. Online SVR [15] is another example. It is a support vector machine learning approach. The resulting algorithm can be applied in real time. Its main advantage is that it can be very responsive when current data draws away for the historical information. However, it needs time for model calibration and high quality training sets.

In order to meet the requirements treated so far, this work proposes four easy implementable adaptive algorithms based on Kalman filter theory. The prediction problem is reformulated as a filtering one by using historical data or the output of a given predictor as pseudo-observations of the future observations. We adopt a Gaussian assumption for the

\* INRIA Rhône-Alpes, NeCS team, Grenoble, France  
luis.leon@inria.fr

\*\* Gipsa-Lab, University Joseph Fourier, CNRS, NeCS team, Grenoble, France  
alain.kibangou@ujf-grenoble.fr

\*\*\* Gipsa-Lab, CNRS, NeCS team, Grenoble, France  
carlos.canudas-de-wit@gipsa-lab.fr

distribution of the error between the unknown current data and the pseudo-observations. The optimality of the filtering being related to the a priori knowledge of the variance of the modeling error, we resort to an adaptive Kalman filtering approach to accommodate such lack of a priori knowledge. A similar Kalman strategy for traffic flow forecasting was proposed in [31], where the model to be fitted is a Seasonal ARIMA model, and the coefficients of the short-term operator are obtained through an adaptive Kalman filter, and the system's observations are computed from the prediction of the Kalman filter. The main disadvantage seen in this approach is the important amount of past data needed to calibrate the model and the time needed to achieve it. Nevertheless in this work, the author succeeded to obtain an online traffic flow forecasting methodology to later predict travel time.

This paper focuses on the development of a responsive multi-step ahead forecasting methodology that, under an easily on-line implementable algorithm can achieve accurate prediction performance, with no requirement of extensive data calibration nor large historical database.

## II. PROBLEM FORMULATION

Given available measurements  $\{\varphi_{t,D}\}$ ,  $t = 0, 1, \dots, t_0$ , of the traffic flow until time  $t_0$  of the day  $D$  and historical data  $\{\varphi_{t,d}\}$ ,  $t = 0, 1, \dots, T$ ,  $d = 1, 2, \dots, D-1$ . Our aim is to predict the flow at time  $t = t_0 + 1, \dots, t_0 + \Delta T$ . For this purpose, we review first some forecasting methods whose efficiency has been demonstrated in the literature. These methods will be compared to the new approaches suggested in this paper. Then in section III, we introduce our adaptive Kalman filtering approach.

### A. Historical increment flow based predictor

Let us consider a first-order autoregressive model for the traffic flow:

$$\varphi_{t,d} = \varphi_{t-1,d} + \varepsilon_{t,d}, \quad (1)$$

where  $\varepsilon_{t,d}$  is a realization of a white gaussian random variable  $\varepsilon_t \sim \mathcal{N}(\mu_{\varepsilon,t}, \sigma_{\varepsilon,t}^2)$ . It is then obvious that the minimum variance one-step ahead predictor is given by

$$\hat{\varphi}_{t,D} = \varphi_{t-1,D} + \mu_{\varepsilon,t}. \quad (2)$$

The underlying idea of this predictor is to make use of the mean value of the historical increments  $\varepsilon_{t,d}$  in order to achieve the flow prediction at the time  $t$ .

The multi-step ahead predictor counterpart is given by:

$$\hat{\varphi}_{t_0+\tau,D} = \begin{cases} \varphi_{t_0+\tau-1,D} + \mu_{\varepsilon,t_0+\tau} & \text{if } \tau = 1 \\ \hat{\varphi}_{t_0+\tau-1,D} + \mu_{\varepsilon,t_0+\tau} & \text{elsewhere} \end{cases}. \quad (3)$$

### B. Gaussian maximum likelihood (GML) approach

In this approach, the flow  $\varphi_{t,d}$  and the flow increment  $\varepsilon_{t,d} = \varphi_{t,d} - \varphi_{t-1,d}$ , with  $d = 1, 2, \dots, D$ , are assumed to be realizations of normally distributed random variables, i.e.,  $\varepsilon_{t,d} \sim \mathcal{N}(\mu_{\varepsilon,t}, \sigma_{\varepsilon,t}^2)$  and  $\varphi_{t,d} \sim \mathcal{N}(\mu_{\varphi,t}, \sigma_{\varphi,t}^2)$  [5]. By maximizing the likelihood function of the flow level at the next time interval, the following predictor is obtained:

$$\hat{\varphi}_{t,D} = \frac{\sigma_{\varphi,t}^2(\mu_{\varepsilon,t} + \varphi_{t-1,D}) + \sigma_{\varepsilon,t}^2\mu_{\varphi,t}}{\sigma_{\varphi,t}^2 + \sigma_{\varepsilon,t}^2} \quad (4)$$

Means and variances involved in (4) are computed using historical data. The corresponding multi-step ahead predictor is then given by:

$$\hat{\varphi}_{t_0+\tau,D} = \frac{\sigma_{\varphi,t_0+\tau}^2(\gamma_{t_0+\tau-1,D} + \mu_{\varepsilon,t_0+\tau}) + \sigma_{\varepsilon,t_0+\tau}^2\mu_{\varphi,t_0+\tau}}{\sigma_{\varphi,t_0+\tau}^2 + \sigma_{\varepsilon,t_0+\tau}^2} \quad (5)$$

with  $\gamma_{t_0+\tau-1,D} = \varphi_{t_0+\tau-1,D}$  if  $\tau = 1$ ; else  $\gamma_{t_0+\tau-1,D} = \hat{\varphi}_{t_0+\tau-1,D}$ .

### C. Constant and heuristics based predictor

This deterministic methodology is based on a combination between a constant and a nonparametric predictor [21]. The constant predictor acts for small prediction horizon, while with the increasing of the horizon heuristics are used for prediction. Based on this consideration, the following predictor was proposed:

$$\hat{\varphi}_{t_0+\tau,D} = \varphi_{t_0+\tau,D}^H + K\Delta\varphi_{t_0,D}, \quad (6)$$

with

$$\Delta\varphi_{t_0,D} = \varphi_{t_0,D} - \varphi_{t_0,D}^H \quad (7)$$

$$K = \begin{cases} \eta(1 - \frac{\tau}{\Delta t_{\max}}), & \text{if } 0 < \tau \leq \Delta t_{\max} \\ 0, & \text{if } \tau > \Delta t_{\max} \end{cases},$$

$\varphi_{t_0,D}^H$  being the flow level of the averaged historical data,  $\eta$  being the weight of the mismatch between the current and the historical data, and  $\Delta t_{\max}$  the maximum horizon for the constant model. The optimal values for the parameters  $\eta$  and  $\Delta t_{\max}$  will be taken from the results presented in [21], i.e.  $\eta = 0.57$ ,  $\Delta t_{\max} = 37$ .

## III. ADAPTIVE KALMAN FILTERING APPROACH

Let us consider the following Gauss-Markov model:

$$\begin{aligned} x_t &= x_{t-1} + w_t \\ y_t &= H_t x_t + v_t, \end{aligned} \quad (8)$$

where  $w_t \sim \mathcal{N}(q_t, Q_t)$  and  $v_t \sim \mathcal{N}(0, R_t)$  are independent from the initial condition of the state  $x_t$  and mutually independent. The sub-optimal  $\tau$ -step ahead predictor is given by  $\hat{x}_{t+\tau|t} = \hat{x}_t + \sum_{i=1}^{\tau} q_{t+i}$  [32], where  $\hat{x}_t$  stands for the output of a Kalman filter, i.e.

$$\hat{x}_{t|t} = (I - K_t H_t) \hat{x}_{t-1|t-1} + K_t y_t + (I - K_t H_t) q_t, \quad (9)$$

with

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1}, \quad (10)$$

$$P_{t|t} = (I - K_t H_t) P_{t|t-1}, \quad (11)$$

and

$$P_{t|t-1} = P_{t-1|t-1} + Q_{t-1}. \quad (12)$$

Clearly, in such a framework a multi-step ahead Kalman predictor is not well indicated since the prediction consists in keeping constant the future values of the state plus

an additive term given by the mean value of the process noise. Therefore, we suggest to transform a multi-step ahead prediction problem to a filtering one. For this purpose, we suggest considering pseudo-observations. Two kinds of pseudo-observations are considered.

The first one corresponds to the use of the average historical flow as noisy measurements of the current flow during the prediction interval, i.e. using the historical data base as the system's measurements. Since traffic volume can vary substantially depending on various external factors, the major drawback in this strategy is the high reliability on the historical data. Thus, the model might not be quite responsive to eventual structural changes.

The second kind of pseudo-observations is obtained by using the predicted flows of the model proposed in [21] as observations. Since the predictor in [21] does not represent an important burden in the forecasting computation, it can be combined with the Kalman filter. The main motivation to propose this second strategy is the fact that the model in [21] was proven to capture easily special events in the current traffic flow, i.e. significant mismatches between on-line and historical data.

When using these pseudo-observations, the noise statistics are a priori unknown. Therefore, we suggest resorting to an adaptive Kalman filtering approach, where the mean and the variance of both process and observation noises are estimated with the available data. This problem has been widely studied in the literature, and several results have been presented [33]–[36]. Because of its simplicity and its ability to handle both systematic and random errors, in this paper, we resort to the method proposed in [36].

#### A. Online noise statistics estimation

1) *Statistics of the observation noise:* From the observation model, an approximation of the observation noise is given by  $r_t = y_t - H_t x_{t|t-1}$ . By considering  $r_t$  as samples of the observation noise, given  $N$  consecutive observations, for instance from  $t = t_0 - N + 1$  to  $t = t_0$ , an unbiased estimate of the mean is given by

$$\hat{r} = \frac{1}{N} \sum_{t=1}^N r_t. \quad (13)$$

Following the strategy in [36] the unbiased estimation of  $R$  becomes

$$\hat{R} = \frac{1}{N-1} \sum_{t=1}^N ((r_t - \hat{r})(r_t - \hat{r})^T - \frac{(N-1)}{N} H_t P_{t|t-1} H_t^T). \quad (14)$$

2) *Statistics of the process noise:* An approximation of the process noise is given by considering the samples of the state noise  $q_t = x_{t|t} - x_{t-1|t-1}$ , for instance from  $t = t_0 - N + 1$  to  $t = t_0$ . Following the previous, the estimation of the mean and covariance for the state noise statistics are

$$\hat{q} = \frac{1}{N} \sum_{t=1}^N q_t. \quad (15)$$

$$\hat{Q} = \frac{1}{N-1} \sum_{t=1}^N ((q_t - \hat{q})(q_t - \hat{q})^T - \frac{(N-1)}{N} (P_{t-1|t-1} - P_{t|t})). \quad (16)$$

In what follows, we consider two different state-space models giving rise to different prediction methods.

#### B. Filtering based on a first order modeling of the flow (KF1)

Let us assume that two consecutive flow levels,  $\varphi_{t,D}$  and  $\varphi_{t-1,D}$ , are related through a first order model

$$\varphi_{t,D} = \varphi_{t-1,D} + w_{t,D}. \quad (17)$$

The measurement model is

$$y_{t,D} = \varphi_{t,D} + (v_{t,D} - \hat{r}), \quad (18)$$

where  $w_{t,D} \sim \mathcal{N}(q_t, Q_t)$  and  $v_{t,D} \sim \mathcal{N}(r, R)$ . The pseudo-observations  $y_{t,D}$  can be averaged historical flow or predictions using the method in [21]. In the first case the algorithm is named KF1-I while in the second case it is named KF1-II. The process and observation noises are modeled as. Note that by keeping  $R$  constant, we can guarantee, in some sense, that at least for the first steps ahead predictions the trend of the on-line flow is followed.

Given the following sequence of pseudo-observations  $\{y_t\}_{t=t_0-N+1, \dots, t_0+\Delta T}$ , and the actual measurements of the flow  $\{\varphi_{t,D}\}_{t=t_0-N+1, \dots, t_0}$ , the KF1 forecasting method acts as follows:

- Initialization:

- $\hat{\varphi}_{0|0} = \varphi_{0,D}$ ,  $P_{0|0}$  is initialized with a small positive value,
- $\hat{r} = \frac{1}{N} \sum_{t=t_0-N+1}^{t_0} (y_{t,D} - \varphi_{t,D})$ ,
- $R = \frac{1}{N-1} \sum_{t=t_0-N+1}^{t_0} (y_{t,D} - \varphi_{t,D} - \hat{r})^2$ ,
- $\hat{q}_0 = \frac{1}{N} \sum_{t=t_0-N+1}^{t_0} (\varphi_{t,D} - \varphi_{t-1,D})$ ,
- $\hat{Q}_0 = \frac{1}{N-1} \sum_{t=t_0-N+1}^{t_0} (\varphi_{t,D} - \varphi_{t-1,D} - \hat{q}_0)^2$ .

- Prediction process: for  $t = t_0 + 1, \dots, t_0 + \Delta T$

- Apply the Kalman filtering equations (9)-(12) to the state-space model (17)-(18), where  $\varphi_{t,D}$  stands for the state  $x_t$ .
- If  $t > t_0 + N$ , update the mean and the variance of the process noise using (15) and (16).

The choice of the number  $N$  of past observations is related to how far into the past we have to look at to guarantee that any structural change in the current flow is taken into account. Notice that in (10)  $R_t = R$ .

#### C. Filtering based on a higher-order model (KF2)

Let us assume the following higher order model for the traffic flow  $\varphi_{t,D}$ :

$$\begin{aligned} \varphi_{t,D} &= \varphi_{t-1,D} \theta^1_{t,D} + \dots + \varphi_{t-n,D} \theta^n_{t,D} + (v_{t,D} - \hat{r}_t) \\ \varphi_{t,D} &= \Psi^T_{t,D} \Theta_{t,D} + (v_{t,D} - \hat{r}_t) \end{aligned} \quad (19)$$

We also assume a random walk model for the regression coefficients vector:

$$\Theta_{t,D} = \Theta_{t-1,D} + (w_{t,D} - \hat{q}_t) \quad (20)$$

The process and observation noises are modeled as  $w_{t,D} \sim \mathcal{N}(q_t, Q_t)$  and  $v_{t,D} \sim \mathcal{N}(r_t, R_t)$ . In contrast to the KF1 approach, here the state  $x_t$  is constituted with the coefficients of the regression vector  $\Theta_{t,D}$ . Therefore, the aim of the Kalman filter is to provide an estimate  $\hat{\Theta}_{t,D}$  of the adequate coefficients  $\Theta_{t,D}$ . Then the multi-step ahead flow prediction is computed as

$$\hat{\phi}_{t+\tau,D} = \Psi_{t+\tau,D}^T \hat{\Theta}_{t+\tau-1,D}, \quad (21)$$

where

$$\begin{aligned} \Psi_{t+\tau,D} &= (\varphi_{t,D}^c, \dots, \varphi_{t-n+1,D}^c) \text{ if } \tau = 1 \\ \Psi_{t+\tau,D} &= (\varphi_{t+\tau-1,D}^h, \dots, \varphi_{t+1,D}^h, \varphi_{t,D}^c, \dots, \varphi_{t+\tau-n,D}^c) \\ &\text{if } \tau \in \{2, \dots, n\} \\ \Psi_{t+\tau,D} &= (\varphi_{t+\tau-1,D}^h, \dots, \varphi_{t+\tau-n,D}^h) \text{ if } \tau \in \{n+1, \dots, \Delta\} \end{aligned}$$

$\varphi_{t,D}^c$  and  $\varphi_{t,D}^h$  represents the online measurements and pseudo-observations based on historical flow levels at time  $t$  respectively.

This approach is similar to those in [18]–[20], [30], [37] designed for one-step ahead predictors. The multi-step ahead predictors are obtained by using, as previously, pseudo-observations from averaged historical flow (KF2-I) or predictions from method in [21] (KF2-II). The proposed KF2-I and KF2-II methods act as follows:

- Initialization:
  - $\hat{\theta}_{0|0}$  is initialized by solving a linear regression problem with the available current flow measurements from  $t = t_o - N$  to  $t = t_o - 1$ .
  - The error covariance matrix  $P_{0|0}$  is accordingly initialized with a small positive value.
  - $\hat{r}_o = \frac{1}{N} \sum_{t=t_o-N}^{t_o-1} (\phi_{t,D} - \Psi_{t,D} \Theta_{t,D})$ ,
  - $\hat{R}_o = \frac{1}{N-1} \sum_{t=t_o-N}^{t_o-1} (\phi_{t,D} - \Psi_{t,D} \Theta_{t,D} - \hat{r}_o)^2$ .
  - $\hat{q}_o = \frac{1}{N} \sum_{t=t_o-N}^{t_o-1} (\Theta_{t,D} - \Theta_{t-1,D})$ ,
  - $\hat{Q}_o = \frac{1}{N-1} \sum_{t=t_o-N}^{t_o-1} (\Theta_{t,D} - \Theta_{t-1,D} - \hat{q}_o)^2$ .
- Prediction process: for  $t = t_0, \dots, t_0 + \Delta T - 1$ 
  - Apply the Kalman filtering equations (9)-(12) to the state-space model (19)-(20).
  - If  $t > t_0 + N$ , update the mean and the covariance matrix of the process and state noise using (13)-(16).
  - Predict the flow using (21)

## IV. EXPERIMENTAL RESULTS

### A. Data description and test design

The traffic volume data used to evaluate the proposed methods were obtained using actual measurements of the

Grenoble south ring aggregated into 1-minute interval on a time window of 24 hours. A database was created, by adding white noise and filtering the resulting data with low-pass filters whose parameters were randomly generated. We obtained 200 flow profiles a half of these profiles were used for heuristics and the other half for validation.

Based on these data, the performance of seven predictors are compared. As it was aforementioned, all of them depend on the historical average data. Thus, it is assumed that the current flow profile belongs to a certain cluster. Two commonly used criteria are employed to evaluate and compare the predictor performance. The mean absolute percentage error (MAPE) and the absolute percent error (APE):

$$MAPE(\%) = \frac{1}{N} \sum_{t=1}^N \left| \frac{\hat{\phi}_{t,D} - \phi_{t,D}}{\phi_{t,D}} \right| \times 100\% \quad (22)$$

$$APE_t(\%) = \frac{|\hat{\phi}_{t,D} - \phi_{t,D}|}{\phi_{t,D}} \times 100\% \quad (23)$$

where  $\phi_i$  and  $\hat{\phi}_i$  stand for the true and the predicted value of the flow at point  $t$  respectively. The prediction performance in tables (I)-(II), is assessed through mean and standard deviation of the MAPE obtained for each predictor in the 100 runs.

### B. Test results and analysis

After testing different choices of the parameters of the filters we selected,  $N = 4$  for the online noise statistics estimate, since it have the minimal APE, and  $p = 4$  as order of the state for KF2.

Given the level of congestion of the considered demand flow data-set, three scenarios were carried out, at 9:00 am, 19:00, and 17:00. The assessment of the predictors for the first two scenarios was achieved by considering three prediction horizons, 15, 30, and 45 minutes. For the third case (17:00) only a horizon of 45 minutes will be evaluated.

1) *Case 1:* Table I presents the forecasting accuracy for the seven predictors at 9:00 am. Several points are worth pointing out. The first is the fact that both Kalman filter based predictors using the average historical flow (KF1-I and KF2-I) are clearly outperformed by the predictors using outputs of the predictor [21] as pseudo-observations (KF1-II and KF2-II). The second thing is that for the shortest prediction horizon, GML approach gives the best results, following Kalman filter approaches 1-II and 2-II respectively. However, as the prognosis horizon increases, KF1-II and KF2-II present better results. It is also important to say that the predictor [21] (Const. and heur.) exhibited comparable results, but slightly inferior, to the ones using a mixture between this and Kalman filter strategy. This is actually an interesting result, since as said previously, our objective was to find a predictor easily implementable online with low computational burden and high accuracy.

2) *Case 2:* Table II, presents the forecasting accuracy for the seven methodologies under study at 19:00. As the case before, the Kalman filter based predictors 1-II and 2-II clearly outperform KF1-I and KF2-I. However, in contrast with the

previous example, at 15 minutes of prediction the GML and historical average flow predictors have better results when considering mean value of the MAPE, nevertheless they exhibit higher standard deviation in the performance. This fact is due to the eventual difference between certain flow profiles, among the 100 used for the model validation, and the historical data base.

It is also seen that for all horizons the GML-based predictor outperforms all predictors in terms of mean value, when considering also standard deviation however, KF 1-II have in average better results, even better than [21]. Taking all this into account, this scenario clearly shows that the proposed KF approaches exhibit higher robustness.

3) *Case 3:* In order to present a graphic idea of the prediction results, let us take an example of a forecasting starting at 17. Figure 1 shows both actual and mean historical flow profile along 45 minutes horizon. We can see that the averaged historical flow lags behind the current flow.

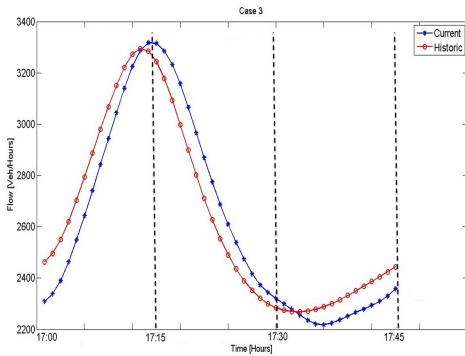


Fig. 1. Historical flow average and current flow. Time between 17:00 and 17:45

TABLE I  
FORECASTING PERFORMANCE-@9:00 AM

Predictors @9:00am	15 minutes	30 minutes	45 minutes
	MAPE(%)	MAPE(%)	MAPE(%)
KF1-I	$2.28 \pm 1.50$	$2.23 \pm 1.11$	$2.24 \pm 0.95$
KF1-II	$1.61 \pm 1.13$	$1.87 \pm 0.96$	$1.96 \pm 0.90$
KF2-I	$2.74 \pm 1.97$	$2.55 \pm 1.30$	$2.40 \pm 1.02$
KF2-II	$1.72 \pm 1.17$	$1.92 \pm 0.96$	$1.99 \pm 0.88$
Const. and heur.	$1.73 \pm 1.17$	$1.93 \pm 0.97$	$2.00 \pm 0.90$
GML	$1.52 \pm 1.33$	$2.28 \pm 1.49$	$3.00 \pm 1.90$
Hist. inc. flow	$1.78 \pm 1.40$	$2.99 \pm 2.42$	$4.55 \pm 4.37$

TABLE II  
FORECASTING PERFORMANCE-@19:00

Predictors @19:00	15 minutes	30 minutes	45 minutes
	MAPE(%)	MAPE(%)	MAPE(%)
KF1-I	$4.01 \pm 2.33$	$4.36 \pm 2.12$	$4.46 \pm 1.96$
KF1-II	$2.87 \pm 1.66$	$3.50 \pm 1.66$	$3.89 \pm 1.64$
KF2-I	$4.78 \pm 2.84$	$4.88 \pm 2.39$	$4.84 \pm 2.22$
KF2-II	$3.07 \pm 1.71$	$3.69 \pm 1.75$	$4.03 \pm 1.77$
Const. and heur.	$3.00 \pm 1.68$	$3.65 \pm 1.74$	$4.01 \pm 1.78$
GML	$2.56 \pm 3.50$	$3.40 \pm 4.98$	$3.87 \pm 4.02$
Hist. inc. flow	$2.76 \pm 3.57$	$3.83 \pm 5.25$	$4.60 \pm 5.48$

Figure 2 depicts the probability of each forecasting method to be equal or lower to the value of APE shown in the horizontal axis in a 45 minutes prediction horizon. It is clear the superiority of the method KF2-II over the other ones, being the second best KF2-I. While in overall the predictor proposed by [21] performed similar to KF1-I and KF1-II. It is important to point out that for the specific scenario considered the GML-based predictor had the lowest performance among all the predictors. To have an idea, methods KF2-I and KF2-II have a 90% probability to present a value of APE lower or equal to 4%, whereas predictor [21] has a 65% and GML-based predictor 15% approximately. This is an important fact since this shows that our forecasting methods can be more responsive to eventual structural changes in the current traffic flow. Specially the approach KF2-II, that performs even better than the one proposed by [21].

It is seen hence that, firstly, as pointed out in [15], GML approach is highly dependant on the quality of the database and its closeness to the online flow. Secondly, even if KF1-II (pseudo-observations obtained from [21]) depends also on the database available, it presents more robust forecasting results when the historical and the current flows do not match precisely.

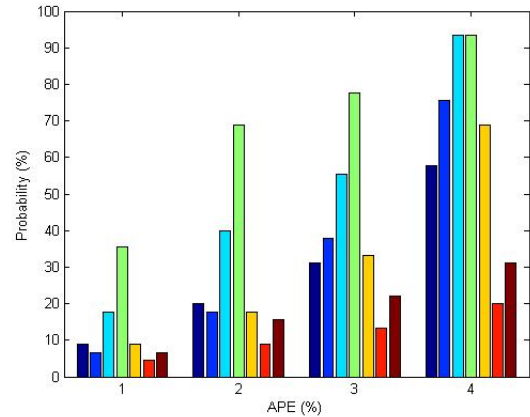


Fig. 2. Probability for achieving an APE lower or equal than the value in the horizontal axis for different forecasting algorithms organized from left to right as KF1-I, KF1-II, KF2-I, KF2-II, Constant and heuristics, GML, and Historical increment flow). To=17:00 with a horizon of 45 minutes

## V. CONCLUSIONS

In this paper, we have proposed four adaptive Kalman filter approaches for the multi-step ahead forecasting of short-term traffic flow and compared their performances to other prediction algorithms, such as [5], [21]. It was proved that two of the Kalman based predictors, the ones using the result of prediction given by [21] to compute the innovations, performed well in all scenarios tested, sometimes even better than [5] and certainly better than [21]. Since Kalman filter and the constant and heuristics based model proposed in [21] are very easy to be implemented on-line, do not represent a heavy computational burden, and need a low calibration

effort, their combination as proposed herein offer very interesting alternatives for improving online traffic forecasting.

Extension of the work presented herein may address the estimation of other traffic parameter such as travel time, using forecasting bounds defined by the Kalman covariance matrix.

#### ACKNOWLEDGMENT

The research leading to these results has received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement no 257462 HYCON2 Network of excellence, and the MOCOPo project.

#### REFERENCES

- [1] D. Pisarski and C. Canudas-de Wit, "Analysis and design of equilibrium points for the cell-transmission traffic model," in *Proc. of IEEE American Control Conference*, Montreal, Canada, June 2012, pp. 5763–5768.
- [2] —, "Optimal balancing of road traffic density distributions for the cell transmission model," in *Proc. of IEEE Conference on Decision and Control*, Maui, Hawaii, USA, December 2012, pp. 6969–6974.
- [3] B. Van Arem, H. Kirby, M. Van Der Vlist, and J. Whittaker, "Recent advances and applications in the filed of short-term traffic forecasting," *International Journal of Forecasting*, vol. 13, pp. 1–12, 1997.
- [4] E. Vlahogiannia, J. Golias, and M. Karlaftisa, "Short term traffic forecasting: overview of objectives and methods," *Transport Reviews*, vol. 24, no. 5, pp. 553–557, September 2004.
- [5] W. Lin, "A Gaussian maximum likelihood formulation for short-term forecasting of traffic flow," in *IEEE Intelligent Transportation Systems Conference Proceedings*, Oakland, CA, USA, August 2001.
- [6] B. Smith, B. Williams, and K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 4, pp. 303–321, 2002.
- [7] M. Der Voort, M. Dougherty, and S. Watson, "Combining Kohonen maps with ARIMA time series models to forecast traffic flow," *Transportation Research Part C*, vol. 4, no. 5, pp. 307–318, 1996.
- [8] B. Williams, K. Prya, and D. Brown, "Urban freeway traffic flow prediction. application of seasonal autoregressive integrated moving average and exponential smoothing models," *Transportation Research Record*, vol. 1644, pp. 179–188, 1998.
- [9] B. Smith and M. Demetsky, "Multiple-interval freeway traffic flow forecasting," *Transportation Research Record*, vol. 1544, pp. 136–141, 1996.
- [10] B. Smith, B. Williams, and K. Oswald, "Parametric and nonparametric traffic volume forecasting," in *Proceedings of the 80th TRB Annual Meeting*, Washington, DC, USA, 2000.
- [11] G. Davis and N. Niham, "Adaptive forecasting of freeway traffic congestion," *Transportation Research Record*, vol. 1287, pp. 29–33, 1991.
- [12] S. Haykin, *NN: A comprehensive Foundation*. NY: Macmillan, 1994.
- [13] S. Clark, M. Dougherty, and H. Kirby, "The use of neural networks and time series models for short-term traffic forecasting: a comparative study," in *Proceedings of the 21st summer Annual Meeting*, 1993.
- [14] S. Ishak and C. Alecsandru, "Optimizing traffic prediction performance of neural networks under various topological, input and traffic condition," in *Proceedings of the TRB 82nd Annual Meeting*, Washington, DC, USA, 2003.
- [15] M. Castro-Neto, Y. Jeong, M. Jeong, and L. Hana, "Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions," *Journal Expert Systems with Applications: An International Journal archive*, vol. 36, no. 3, pp. 6164–6173, April 2009.
- [16] B. Smith and M. Demetsky, "Traffic flow forecasting: comparison of modelling approaches," *Journal of Transportation Engineering*, vol. 123, no. 4, pp. 261–266, 1997.
- [17] B. Williams, P. Durvasaula, and D. Brown, "Urban traffic traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models," *Transportation Research Record*, vol. 1644, pp. 132–144, 1998.
- [18] X. Yuanchang, Z. Yunlong, and Y. Zhirui, "Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition," *Computer-Aided Civil and Infrastructure Engineering*, vol. 22, no. 5, pp. 326–334, July 2007.
- [19] A. Stathopoulos and A. Karlaftis, "A multivariate state-space approach for urban traffic flow modelling and prediction," *Transportation Research Part C*, vol. 11, no. 2, pp. 121–135, April 2003.
- [20] I. Okutani and Y. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transportation Research Part B*, vol. 18, no. 1, pp. 1–11, February 1984.
- [21] R. Chrobok, J. Kaumann, and M. Schreckenberg, "Different methods of traffic forecast based on real data," *European Journal of Operational Research*, vol. 155, pp. 558–568, 2004.
- [22] M. Cetin and G. Comert, "Short-term traffic flow prediction with regime switching models," *Transportation Research Record*, vol. 1965, pp. 23–31, 2006.
- [23] Y. Tang, W. Lam, and P. Ng, "Comparison of four modeling techniques short-term AADTT forecasting in Hong Kong," *Journal of Transportation Engineering*, vol. 129, no. 3, pp. 271–277, 2003.
- [24] W. Lam, Y. Tang, K. Chan, and M. Tam, "short-term traffic flow forecast using Hong Kong Annual Traffic Census," *Transportation*, vol. 33, pp. 291–310, 2006.
- [25] M. Dougherty and M. Cobbet, "Short-term inter-urban traffic forecasts using neural networks," *International Journal of Forecasting*, vol. 13, no. 1, pp. 21–31, 1997.
- [26] B. Abdulhai, H. Porwal, and W. Recker, "Short term freeway traffic flow prediction using genetically-optimized time-delay-based neural networks," *UCB, UCB-ITS-PWP-99-1 (berkeley, CA: Institute of Transportation Studies, University of California, Berkeley)*, 1999.
- [27] S. Innamaa, "Short-term prediction of traffic situation using MLP-neural networks," in *Proceedings of the 7th World congress on Intelligent Transportation Systems*, Turin, Italy, November 2000.
- [28] H. Kirby, M. Dougherty, and S. Watson, "Should we use neural networks or statistical models for short-term motorway forecasting," *International Journal of Forecasting*, vol. 13, no. 1, pp. 43–50, 1997.
- [29] M. Danech-Pajouh and M. Aron, "ATHENA: a method for short-term inter-urban motorway traffic forecasting," *Recherche Transport Securite*, vol. 6, pp. 11–16, 1991.
- [30] J. Whittaker, S. Garside, and K. Lindveld, "Tracking and predicting a network traffic process," *International Journal of Forecasting*, vol. 13, no. 1, pp. 51–61, March 1997.
- [31] J. Xia, *Dynamic freeway travel time prediction using single loop detector and incident data*. Kentucky (USA): University of Kentucky Doctoral Dissertations, 2006.
- [32] R. Kalman and R. Bucy, "New results in linear filtering and prediction theory," *Transactions of the ASME. Series D, Journal of Basic Engineering*, vol. 83D, pp. 95–108, 1961.
- [33] A. Sage and G. Husa, "Adaptive filtering with unknown prior statistics," in *Proceedings of the joint Automatic Control Conference*, Colorado, USA, August 1969, pp. 760–769.
- [34] R. Mehra, "Adaptive filtering with unknown prior statistics," in *Proceedings of the joint Automatic Control Conference*, Colorado, USA, August 1969, pp. 760–769.
- [35] A. Sage and J. Melsa, *System identification*. New York (USA): Academic press, 1971.
- [36] K. Myers and B. Tapley, "Adaptive sequential estimation with unknown noise statistics," *IEEE Transactions on Automatic Control*, vol. 21, no. 24, pp. 520–523, August 1976.
- [37] C. Lan, "A recursive traffic flow predictor based on dynamic generalized linear model framework," in *IEEE Intelligent Transportation Systems Conference Proceedings*, Oakland, California, USA, August 2001, pp. 410–415.