

## Research Article

# Bayesian Dynamic Linear Model with Adaptive Parameter Estimation for Short-Term Travel Speed Prediction

Tai-Yu Ma <sup>1</sup> and Yoann Pigné <sup>2</sup>

<sup>1</sup>Luxembourg Institute of Socio-Economic Research (LISER), 11 Porte des Sciences, 4366 Esch-Sur-Alzette, Luxembourg

<sup>2</sup>Normandie Univ, UNIHAVRE, UNIROUEN, INSA Rouen, LITIS, 76600 Le Havre, France

Correspondence should be addressed to Tai-Yu Ma; tai-yu.ma@liser.lu

Received 27 November 2018; Accepted 5 March 2019; Published 23 June 2019

Academic Editor: Emanuele Crisostomi

Copyright © 2019 Tai-Yu Ma and Yoann Pigné. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Bayesian dynamic linear model is a promising method for time series data analysis and short-term forecasting. One research issue concerns how the predictive model adapts to changes in the system, especially when shocks impact system behavior. In this study, we propose an adaptive dynamic linear model to adaptively update model parameters for online system state prediction. The proposed method is an automatic approach based on the feedback of prediction errors at each time slot without the needs of external intervention. The experimental study on short-term travel speed prediction shows that the proposed method can significantly reduce the prediction errors of the traditional dynamic linear model and outperform two state-of-the-art methods in the case of major system behavior changes.

## 1. Introduction

Accurate short-term traffic prediction plays an important role for successful traffic information system application such as en-route navigation system, traffic control, and traffic congestion management [1]. During past decades, different prediction methods have been proposed to predict traffic states for developing effective traveler information system and real-time traffic management. The methodology for short-term travel time/traffic flow prediction can be classified into the data-driven approach and model-driven approach [2–5]. The model-driven approach consists in applying traffic flow theory to inference traffic state dynamics based on partial observation of traffic data [6–10]. The advantage of the model-driven approach is that it can obtain accurate traffic state estimation with fewer observations. However, the performance of the model-driven approach can be poor if the applied models are not well calibrated [4]. As regards the data-driven approach, it relies on the spatial-temporal correlation of traffic states for which future traffic states can be estimated based on historical time series data. Among different data-driven approaches, which are the focus

of this study, machine learning methods are widely used for traffic characteristics prediction, e.g., neural networks [11, 12], autoregressive integrated moving average models (ARIMA) [13, 14], support vector machine methods [15], nearest neighbor classification methods [16, 17], ensemble learning approach [12], and Bayesian dynamic linear models (DLM)/state space models [18–22], among many others. One of the main issues in short-term traffic prediction is how to dynamically adapt a predicting model to the uncertainty of system behavior changes, in particular in case of accident or unforeseen events. With recent vehicular communication advances in real-time traffic data collection, the development of adaptive short-term traffic prediction methods become an active research area in transportation science and in developing applications based on vehicular communication technology.

In this perspective, the DLM approach provides a systematic approach based on Bayes' theorem for system states updating and prediction. This approach considers system states of interest as unknown stochastic variables to be estimated. The prior distribution of system states is quantified based on historical data. By collecting new data over time, the

posterior distribution of system states can be estimated based on the Bayes' theorem. This sequential learning framework provides an adaptive learning process for handling time series data prediction. It has been shown that model parameters need to be adaptive with system behavior [23]. Fei et al. [19] proposed a DLM for real-time short-term freeway travel time prediction. The model adjusts the variances of disturbance under a user-defined threshold based on the adaptive control theory. However, such an adjustment mechanism is not optimized and relies on the intervention of expert knowledge, raising issues in its generalization in different areas. For this issue, Fei et al. [18] incorporated a Markov switching process in the DLM based on the three-phase traffic flow theory. They showed the Markov switching DLM approach outperforms the ARIMA method.

Another adaptive modeling approach consists in developing methodology to detect change points of system states and update system parameters to catch system behavior changes [24–29]. The change-point detection methods can be designed to monitor prediction errors and detect accidents, providing feedback to adjust model prediction and reduce prediction errors. Comert and Bezuglov [25] applied the hidden Markov model (HMM) and the expectation-maximization (EM) algorithm as a change-point detection method to update the estimation of parameters (i.e., process mean) used in the autoregressive integrated moving average (ARIMA) model. Moreira-Matias and Alesiani [28] proposed a change detection method based on Page-Hinkley change-point detection method [30] for triggering an accident alarm. The threshold for alarm triggering is a user-defined deterministic parameter, corresponding to a tolerable false alarm rate.

In this study, a new online adaptive parameter estimation approach is proposed under the DLM framework to achieve better accuracy of prediction when some external events or system regime changes occur. This method is based on continuously monitoring prediction errors for adaptively adjusting model parameters. The main contribution resides on the adaptive model parameter adjustment design to improve classical DLM approach when unpredicted system behavior changes are detected. The performance of the proposed approach is tested on a simulated road network under accidental scenarios. The performance of the proposed method is compared with classical DLM methods and the benchmark methods, i.e., ARIMA method [31] and Holt-Winters Exponential Smoothing method [32]. Note that we do not intend to extensively compare the proposed approach with other methods, but instead to demonstrate the effectiveness of the proposed method in improving the parameter setting issues of classical DLM approaches.

The rest of the paper is organized as follows. In Section 2, we present the general DLM forecasting framework and different DLM specifications for time series data analysis. In Section 3, a new adaptive parameter estimation method is proposed for online parameter learning to reduce prediction error. Section 4 reports the numerical study on real-time short-term road travel speed prediction under accidental

scenarios. A comparative study with two other state-of-the-art methods is provided. Finally, conclusions are drawn and future extensions are discussed.

## 2. Bayesian Dynamic Linear Model for Traffic State Prediction

A general DLM can be described by an observation equation and a system state equation to model the process of a system [23]. The state equation describes system state evolution mapping from a priori distribution at  $t-1$  to posterior distribution at time  $t$ . The observation equation describes observed measurements at time  $t$  in relation to system states. The evolution of system states over time is assumed to follow a stochastic process with Gaussian errors. A DLM can be written as [23]

$$\mathbf{x}_t = \mathbf{G}_t \mathbf{x}_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim N(\mathbf{0}, \mathbf{W}_t) \quad (\text{state equation}) \quad (1)$$

$$\mathbf{y}_t = \mathbf{F}_t \mathbf{x}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim N(\mathbf{0}, \mathbf{V}_t) \quad (\text{observation equation}) \quad (2)$$

In (1),  $\mathbf{x}_t$  is the system state at time  $t$ .  $\mathbf{G}_t$  is the evolution matrix of  $\mathbf{x}_t$ . In (2),  $\mathbf{y}_t$  is observation at time  $t$ .  $\mathbf{F}_t$  is the design matrix of  $\mathbf{x}_t$ .  $\mathbf{w}_t$  and  $\mathbf{v}_t$  are white noise error terms following normal distribution with 0 mean and variance  $\mathbf{W}_t$  and  $\mathbf{V}_t$ , respectively. It is assumed that  $\mathbf{v}_t$  and  $\mathbf{w}_t$  are mutually independent, i.e.,  $\text{cov}(\mathbf{v}_t, \mathbf{w}_t) = 0$  for  $t = 1, \dots, T$ . The ratio  $\mathbf{w}_t/\mathbf{v}_t$  is called the *signal-to-noise ratio* at time  $t$ . It represents the ratio of system prediction errors  $\mathbf{w}_t$  and observation errors  $\mathbf{v}_t$ . In general,  $\mathbf{W}_t$  and  $\mathbf{V}_t$  are unknown and need to be estimated from data. The general DLM can be represented by a quadruple  $\{\mathbf{F}_t, \mathbf{G}_t, \mathbf{V}_t, \mathbf{W}_t\}$  over time  $t = 1, 2, \dots, T$ . The DLM provides a probabilistic linkage to update the posterior distribution of system states based on a priori distribution and newly available observations over time based on the Bayesian forecasting framework [19, 23, 33]. The Bayesian forecasting framework in the context of traffic speed prediction on a road network is described as follows.

### 2.1. Bayesian Forecasting Framework

*Step 1* (initialization). Initialize system state variables  $\mathbf{x}_0$  (i.e., travel speed on a road section/link) at  $t = 0$ .

$$(\mathbf{x}_0 | \mathbf{D}_0) \sim N(\mathbf{m}_0, \mathbf{C}_0), \quad (3)$$

where  $\mathbf{m}_0$  denotes the estimated means of link travel speed at  $t = 0$ ;  $\mathbf{C}_0$  is the estimated variance based on the initial information set  $\mathbf{D}_0$  (i.e., historical travel speed data on network). Set  $t = t_0$ .

*Step 2* (prior distribution estimation). Estimate the prior distribution of  $\mathbf{x}_t$  as

$$(\mathbf{x}_t | \mathbf{D}_{t-1}) \sim N(\mathbf{a}_t, \mathbf{R}_t) \quad (4)$$

with  $\mathbf{a}_t = \mathbf{G}_t \mathbf{m}_{t-1}$  and  $\mathbf{R}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t' + \mathbf{W}_t$

where  $N(\mathbf{a}_t, \mathbf{R}_t)$  is the normal distribution.  $\mathbf{a}_t$  is the estimated mean of system states, and  $\mathbf{R}_t$  is the estimated variance of

system states. We can observe that increasing  $\mathbf{G}_t$  or  $\mathbf{W}_t$  will amplify the variance of  $\mathbf{x}_t$ .

*Step 3 (one-step forecast).* Estimate one-step forecast for  $\mathbf{y}_t$  as

$$(\mathbf{y}_t | \mathbf{D}_{t-1}) \sim N(\mathbf{f}_t, \mathbf{Q}_t) \quad (5)$$

with  $\mathbf{f}_t = \mathbf{F}_t' \mathbf{a}_t$  and  $\mathbf{Q}_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + \mathbf{V}_t$

where  $\mathbf{f}_t$  is the mean of prediction at  $t$  and  $\mathbf{Q}_t$  is the variance of prediction at  $t$ . We can observed that if  $\mathbf{F}_t$  (i.e., design matrix) is constant,  $\mathbf{f}_t \propto \mathbf{a}_t = \mathbf{G}_t \mathbf{m}_{t-1}$  and  $\mathbf{Q}_t \propto \mathbf{R}_t + \mathbf{V}_t$ .

*Step 4 (posterior distribution at  $t$ ).* Calculate the posterior distribution  $(\mathbf{x}_t | \mathbf{D}_t)$  as

$$(\mathbf{x}_t | \mathbf{D}_t) \sim N(\mathbf{m}_t, \mathbf{C}_t) \quad (6)$$

with  $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t \mathbf{e}_t$  and  $\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t \mathbf{Q}_t \mathbf{A}_t'$

where  $\mathbf{A}_t = \mathbf{R}_t \mathbf{F}_t \mathbf{Q}_t^{-1}$  and  $\mathbf{e}_t = \mathbf{y}_t - \mathbf{f}_t$ . If  $\mathbf{F}_t = \mathbf{1}$ , then  $\mathbf{A}_t = \mathbf{R}_t / \mathbf{Q}_t$ .

*Step 5 (iterate).* Set  $t = t + 1$ . If  $t = T$  then stop; otherwise go to Step 1.

The proof of the one-step forecast and the posterior distribution can be found in [23]. Note that, for the univariate DLM, it has been shown that the covariance of system evolution  $\mathbf{W}_t$  needs to adapt to drastic system behavior changes or regime shift [23]. Regardless of this issue will make a serious prediction bias [12]. However, none of the existing studies propose any adaptive parameter estimation for  $\mathbf{W}_t$  to address this issue.

We specify three DLMs, i.e., first-order DLM, cubic spline smoothing DLM, and second-order DLM with increasing complexity based on the above DLM forecasting framework for travel speed prediction on a road network. The aim is to provide the benchmarks to compare with the performance of the proposed adaptive parameter updating DLM. The three DLMs are described as follows.

*(a) First-Order DLM.* This is the basic DLM which incorporates a mean level term and a Gaussian noisy term to describe system state evolution.

$$y_t = \mu_t + \varepsilon_{\text{obs}}, \quad \varepsilon_t \sim N(0, \sigma_{\text{obs}}^2) \quad (7)$$

$$\mu_t = \mu_{t-1} + \varepsilon_{\text{level}}, \quad \varepsilon_{\text{level}} \sim N(0, \sigma_{\text{level}}^2) \quad (8)$$

$$\mathbf{x}_t = \mu_t,$$

$$\mathbf{F} = \mathbf{G} = \mathbf{1}, \quad (9)$$

$$\mathbf{W} = \sigma_{\text{level}}^2,$$

$$\mathbf{V}_t = \sigma_{\text{obs}(t)}^2$$

*(b) Cubic Spline Smoothing DLM.* This model extends the first-order DLM by incorporating a local linear trend. The resulting system of equations is written as follows:

Equation (4) and

$$\mu_t = \mu_{t-1} + \alpha_{t-1} + \varepsilon_{\text{level}}, \quad \varepsilon_{\text{level}} \sim N(0, \sigma_{\text{level}}^2) \quad (10)$$

$$\alpha_t = \alpha_{t-1} + \varepsilon_{\text{trend}}, \quad \varepsilon_{\text{trend}} \sim N(0, \sigma_{\text{trend}}^2) \quad (11)$$

In terms of the quadruples of DLM, it is equivalent to

$$\mathbf{x}_t = \begin{bmatrix} \mu_t \\ \alpha_t \end{bmatrix},$$

$$\mathbf{F} = [1, 0],$$

$$\mathbf{G} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad (12)$$

$$\mathbf{W} = \begin{bmatrix} \sigma_{\text{level}}^2 & 0 \\ 0 & \sigma_{\text{trend}}^2 \end{bmatrix},$$

$$\mathbf{V}_t = \sigma_{\text{obs}(t)}^2$$

*(c) Second-Order DLM.* This model extends the cubic spline smoothing DLM by introducing a second linear trend to model changes of the trend level. The second-order DLM is described as follows:

Equations (4) and (7) and

$$\alpha_t = \alpha_{t-1} + \beta_{t-1} + \varepsilon_{\text{trend1}}, \quad \varepsilon_{\text{trend1}} \sim N(0, \sigma_{\text{trend1}}^2) \quad (13)$$

$$\beta_t = \beta_{t-1} + \varepsilon_{\text{trend2}}, \quad \varepsilon_{\text{trend2}} \sim N(0, \sigma_{\text{trend2}}^2) \quad (14)$$

$$\mathbf{x}_t = \begin{bmatrix} \mu_t \\ \alpha_t \\ \beta_t \end{bmatrix},$$

$$\mathbf{F} = [1, 0, 0],$$

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad (15)$$

$$\mathbf{W} = \begin{bmatrix} \sigma_{\text{level}}^2 & 0 & 0 \\ 0 & \sigma_{\text{trend1}}^2 & 0 \\ 0 & 0 & \sigma_{\text{trend2}}^2 \end{bmatrix},$$

$$\mathbf{V}_t = \sigma_{\text{obs}(t)}^2$$

In our travel speed prediction context,  $y_t$  is observed data of average link travel speed at time  $t$ .  $\mu_t$  is the unknown average speed at time  $t$ .  $\alpha_t$  is the trend of variation of averages.  $\varepsilon_{\text{level}}$  and  $\varepsilon_{\text{trend}}$  are the corresponding error terms, respectively. Note that more complicated DLMs using a higher-order trend component or combining a systematic seasonal variation component and a regression component can also be specified.

### 3. Adaptive Parameter Updating for DLM

*3.1. Adaptive DLM.* We propose an adaptive parameter updating approach based on the first-order DLM. We have two unknown parameters, i.e.,  $\mathbf{V}_t$  and  $\mathbf{W}_t$ , to be estimated. The two parameters determine the predicted system states and influence the accuracy of prediction. To estimate the unknown model parameters, we can construct the likelihood function based on observed data as a function of unknown parameters. The maximum likelihood estimation approach is used to estimate the parameters [34]. The log-likelihood function is written as follows [33, 35]:

$$LL(\boldsymbol{\theta}) = -\frac{1}{2} \sum_{t=1}^T \log |\mathbf{Q}_t| - \frac{1}{2} \sum_{t=1}^T (\mathbf{y}_t - \mathbf{f}_t)' \mathbf{Q}_t^{-1} (\mathbf{y}_t - \mathbf{f}_t) \quad (16)$$

where  $\boldsymbol{\theta}$  denotes the unknown parameters, i.e.,  $\boldsymbol{\theta} = (\mathbf{V}_t, \mathbf{W}_t)$ .  $\mathbf{Q}_t$  and  $\mathbf{f}_t$  are the variances and means of prediction at time  $t$  (see (5)), respectively. The maximum likelihood estimates (MLE) of parameters can then be obtained by solving the following optimization problem

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} LL(\boldsymbol{\theta}) \quad (17)$$

In classical DLMS, the system parameters  $\boldsymbol{\theta}$  are constant regardless system regime changes. Fei et al. [19] proposed an intervention approach by adjusting the model error covariance based on anticipated changes from additional exterior information and/or expert's knowledge. The drawback is expert's adjustment might be trivial and lack a system-wide control based on the feedback of prediction errors.

Different with existing approach, we propose a two-stage algorithm by first estimating initial parameters  $\boldsymbol{\theta}_0$  based on a training data set and then using an online adaptive parameter updating based on the feedback of one-step prediction errors. It is similar to feedback control to optimize the model parameters. The proposed two-stage adaptive parameter updating approach is described as follows.

*3.2. Online Adaptive Parameter Updating Approach.* The proposed approach estimates the model parameters  $(\mathbf{V}_0, \mathbf{W}_0)$  based on historical data and adaptively optimizes its model parameters over time based on one-step model prediction errors. The approach is described as follows.

*Step 1* (initial parameter estimation).

- (i) Estimate  $\widehat{\mathbf{V}}_0$  and  $\widehat{\mathbf{W}}_0$ : given input training data set  $\mathbf{D}$ , compute MLE estimates of  $\widehat{\mathbf{V}}_0$  and  $\widehat{\mathbf{W}}_0$ . Get  $\hat{\mathbf{v}}_0 = \sqrt{\widehat{\mathbf{V}}_0}$ .
- (ii) Optimize  $\mathbf{W}_0$ : given  $\mathbf{v}_0$ , find the optimal signal-to-noise ratio  $\hat{\mathbf{s}}$  (i.e.,  $\mathbf{w}/\mathbf{v}_0$ ) as

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmin}} G(\mathbf{y} | \boldsymbol{\theta}) \quad (18)$$

where  $G(\mathbf{y} | \boldsymbol{\theta})$  is a loss function defined by the root mean square error. The optimal estimates of model

error covariance for the training data set can then be obtained as  $\widehat{\mathbf{W}}_0^* = \hat{\mathbf{s}}^2 \widehat{\mathbf{V}}_0$ .

*Step 2* (online adaptive parameter updating). Set  $\mathbf{V}_t = \widehat{\mathbf{V}}_0$  and  $\mathbf{W}_t = \widehat{\mathbf{W}}_0^*$  and compute one-step forecast  $\hat{y}_t$  and prediction error  $\varepsilon_t = |y_t - \hat{y}_t|$  based on (4)-(6). Given a predefined tolerable threshold  $\tau$ , update  $\mathbf{W}_{t+1}$  as

$$\mathbf{W}_{t+1} = \begin{cases} (\hat{\mathbf{s}}_{t+1})^2 \mathbf{V}_t & \text{if } \varepsilon_t \geq \tau \\ \mathbf{W}_t & \text{otherwise} \end{cases} \quad (19)$$

where  $\hat{\mathbf{s}}_{t+1} = \underset{\mathbf{s}}{\operatorname{argmin}} G(\mathbf{y} | D_t, \boldsymbol{\theta})$ .  $\mathbf{V}_t$  is kept constant. Note that one can obtain  $\hat{\mathbf{s}}_{t+1}$  without difficulty by the golden section search or the line search approach [36].

The online adaptive parameter updating approach is shown in Figure 1.

*3.3. Measure Metrics for Assessing Prediction Accuracy.* To measure the accuracy of prediction, two metrics are applied: (1) Root Mean Square Error (RMSE); (2) Mean Absolute Error (MAE). The first one computes the mean of squared error terms. The second one reports the mean of absolute errors. The definitions are as follows.

(a) Mean absolute error (MAE) measures the average magnitude of prediction errors by taking into account all observation equally.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (20)$$

where  $N$  is the total number of observation.  $y_i$  and  $\hat{y}_i$  are the observation and prediction values of sample  $i$ , respectively.

(b) Root Mean Square Error (RMSE) is a second-order measure for prediction errors.

$$\text{RMSE} = \left[ \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \right]^{1/2} \quad (21)$$

The RMSE is a kind of second-order measure of prediction errors. Note MAE and RMSE provide similar measures for quantifying the model prediction errors. However, RMSE weights more to large errors providing more desired property when large errors are undesirable.

## 4. Experimental Study

*4.1. Experimental Settings and Link Speed Data.* We generate realistic travel speed data by microscopic traffic simulations implemented by SUMO [37], a widely used microscopic traffic simulation. The test area is selected from Luxembourg City and its surrounding areas [38, 39]. The travel speed data is collected from the simulator on 13 main road sections in Luxembourg City (Figure 2) in the morning peak-hour period from 7:00-9:00. Link average travel speed is aggregated in a 5-minute aggregation interval (i.e., 7:00, 7:05, 7:10, ..., 9:00) as is the case for most

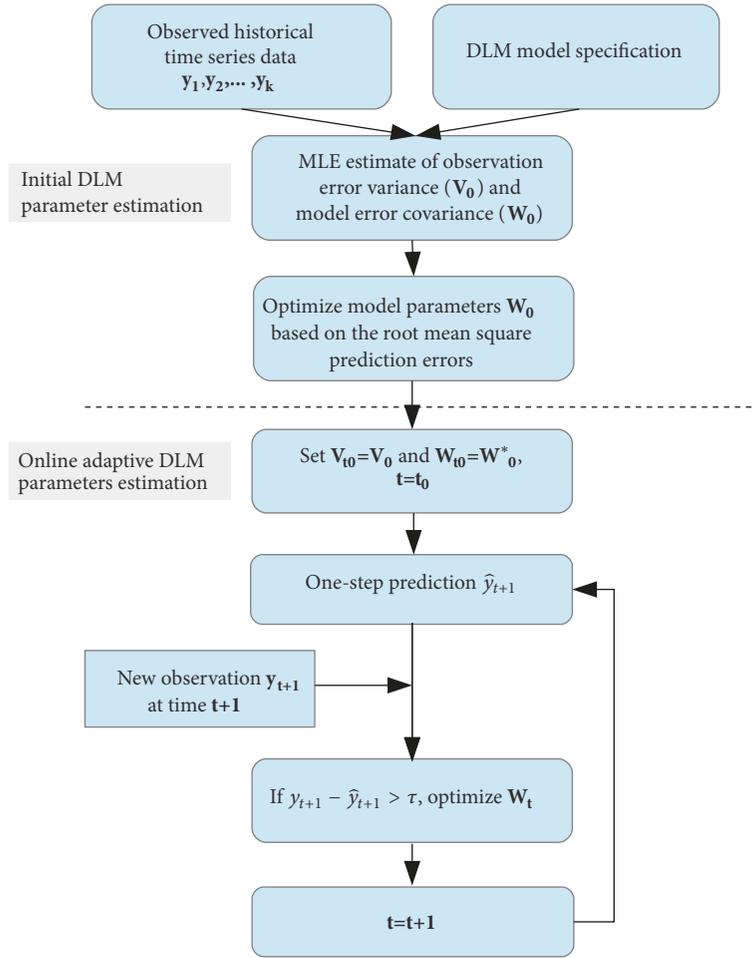


FIGURE 1: Online adaptive parameter updating approach.

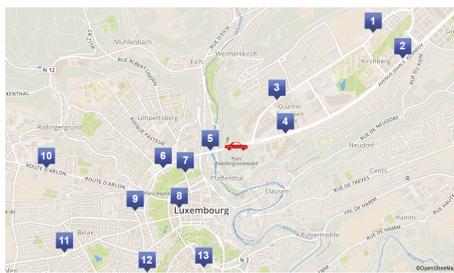


FIGURE 2: Luxembourg road network and traffic count positions. There are totally 13 traffic counts located over main road sections around Luxembourg City and the quarter of European Union Institutions (Kirchberg). The red car is the accident scenario occurring on Grand Duchess Charlotte Bridge standing in way for the direction to east during a.m. 7:30 to a.m. 8:00.

realistic applications [40]. The data is freely available at <https://github.com/pigne/2019-simulations-DLM>.

We consider two scenarios: normal traffic without accidents and traffic with an accident occurred during 7:30-8:00 on the Grand Duchess Charlotte Bridge (see Figure 2)

connecting European Institution quarter and Luxembourg City center. Travel demand is generated based on the realistic LUST traffic demand scenarios for Luxembourg [41] which represents the daily mobility patterns of peoples working/living in the study area. We generate one training data set for initial DLM parameter estimation under normal traffic situation and one test data set under accident situations. The aim is to test the performance of the proposed adaptive parameter updating approach under unforeseen event. The generated traffic patterns are different from one day to another due to stochastic behavior of traffic.

The travel speed profiles on three road sections around the accidental site (road sections 3, 4, and 5) for normal and accidental scenarios are shown in Figure 3. In the normal traffic scenario, there are some frustrations on road section 3W and 4W (direction for Luxembourg City center). When an accident occurs (see the right part in Figure 3), traffic is heavily impacted on the road section 5 for both directions and on the road sections 3 and 4 to its east direction. We can find there is significant travel speed reduction on nearby roads due to the accident event. The numerical is executed by DLM Matlab Toolbox (<https://mjilaine.github.io/dlm/>) using

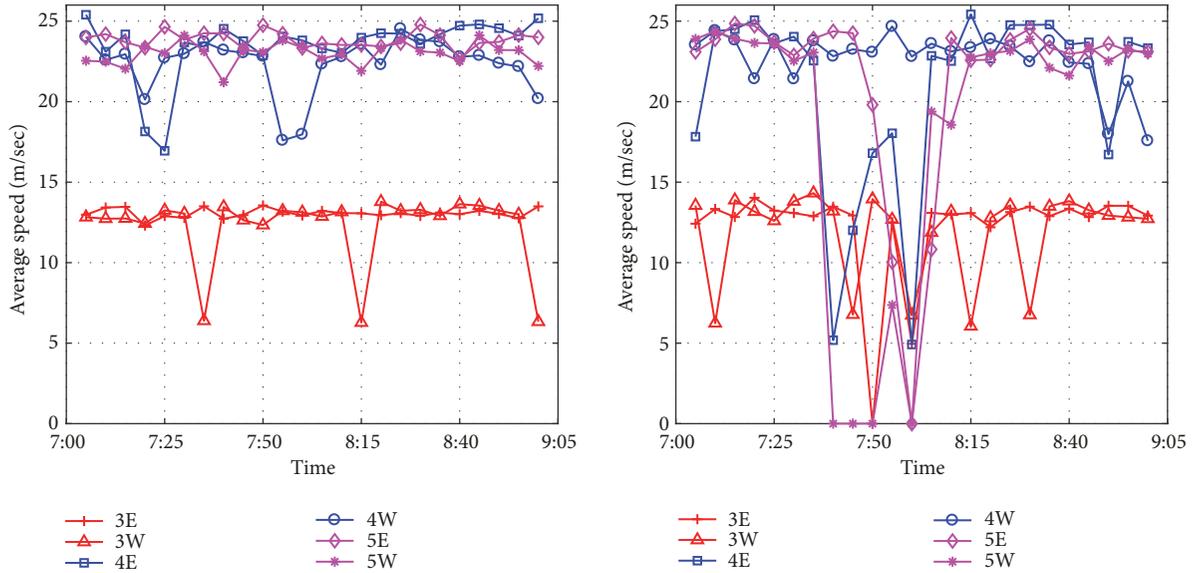


FIGURE 3: Examples of observed average speeds on road sections 3, 4, and 5 for both directions (direction east (E) and direction west (W)). Left: normal traffic scenario. Right: accident scenario (accident on the Grand Duchess Charlotte Bridge).

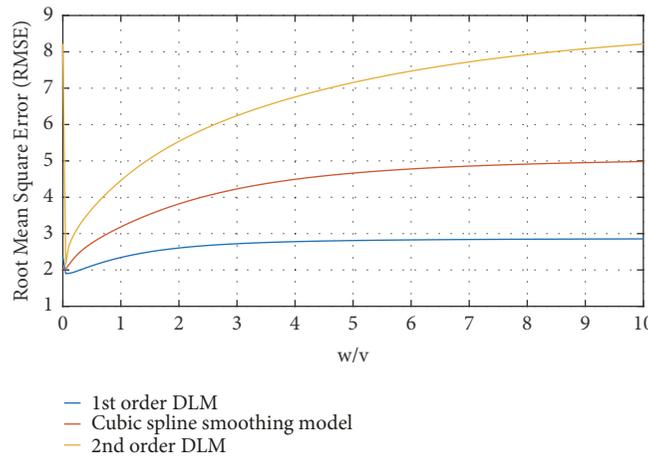


FIGURE 4: Influence of the signal-to-noise ratio ( $w/v$ ) on one-step forecast accuracy of DLMs on traffic data on road section 4E.

a Dell Latitude E5470 laptop with win64 OS, Intel i5-6300U CPU, 2 Cores and 8GB memory.

4.2. Result

4.2.1. Initial Parameter Estimation of the Adaptive DLM. We use the MLE method to obtain an initial estimate of  $V_0$  and  $W_0$  based on the training data set, i.e., observations in the normal traffic scenario. The optimal signal-to-noise ratio with a minimal RMSE value of the one-step forecast can be obtained (Figure 4). The RMSE values is a function of signal-to-noise ratio which decreases at the beginning and then increases until a stable value when increasing the signal-to-noise ratio. We estimate the optimal signal-to-noise ratios and optimal model error covariance for each link. Figure 4 shows the first-order DLM obtains best fits

(i.e., lowest RMSE) compared to the second-order DLM and the cubic spline smoothing DLM model. In normal traffic scenario, traffic speed presents small fluctuation for most of the time. The evolution function in the first-order DLM captures smooth changes of mean state traffic evolution with best goodness-of-fit. However, higher-order DLMs might overfit local trend resulting in higher prediction error.

Figure 5 reports the local trends of the DLMs with and without optimizing signal-to-noise ratio to minimize the RMSE. We found that, after optimizing the signal-to-noise ratio, the fitted Kalman filter smoother becomes more adaptive to observations (on the right side of Figure 5). The DLM Kalman filter smoother has smaller variance with MLE parameters. In terms of one-step forecast accuracy, the prediction accuracy is improved when applying the optimized signal-to-noise ratio in the MLE models.

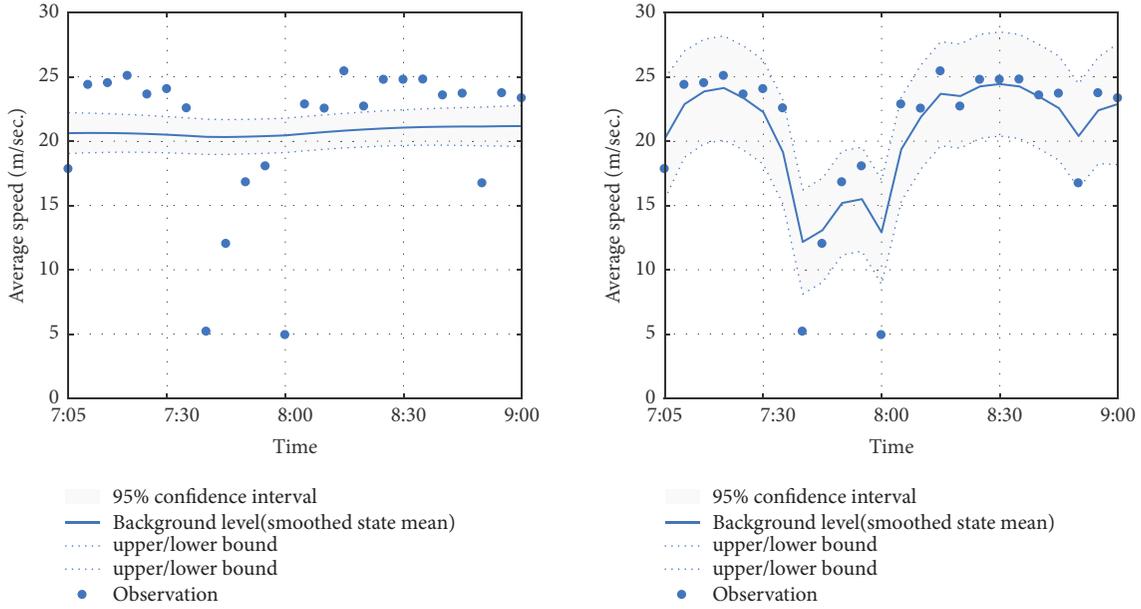


FIGURE 5: First-order DLM Kalman filter smoother fitted on average travel speed on link 4E. On the left: parameters ( $V_0, W_0$ ) estimated by MLE; on the right: after optimization.

TABLE 1: Average prediction accuracy of the adaptive DLM in accident scenarios.

Measure metrics	Road Section	First-order DLM	Cubic spline smoothing DLM	Second-order DLM	AR(2)	HW Exp. Smoothing	One-step shift predictor	Adaptive DLM
RMSE	3, 4 and 5	4.901	5.450	5.537	4.653	4.612	4.830	4.480
	Others	1.948	2.155	2.530	2.587	2.596	2.336	2.020
MAE	3, 4 and 5	3.398	4.029	4.071	2.63	2.666	2.814	2.595
	Others	1.511	1.655	1.913	1.446	1.469	1.741	1.500

**4.2.2. Online Adaptive Parameter Updating.** We test the performance of the proposed approach to the traffic accident scenario. As we can see on Figure 2, when the traffic accident occurs, its upstream and downstream road sections, i.e., road sections 3, 4, and 5, would have significant impacts. Hence it would be interesting to investigate the performance of the proposed method on these road sections.

Table 1 shows the adaptive DLM significantly outperforms the other methods for the cases of major changes in traffic on road sections 3, 4, and 5. The average RMSE of the adaptive DLM over the road sections 3, 4, and 5 is 4.480, compared to the HW Exponential Smoothing method (4.612), AR(2) (4.653) and the three DLM approaches. It outperforms the simple one-step shift predictor (i.e., using observations at time  $t$  as predictors for  $t+1$ ) in both accidental and normal traffic road sections. The values of the MAE measure claim the same conclusion. However, on the other road sections, the adaptive DLM performs similar well compared with the other approaches. The average execution of the adaptive DLM for each road section is 0.1082 second.

To illustrate the effectiveness of reducing prediction errors of the proposed method in case of major changes in traffic, we investigate two road sections which are significantly impacted by the accident, i.e., 4E and 5W. We

can find travel speed quickly drop at about 7:40, and the traffic becomes fluid at about 8:05 on both road sections (see Figure 3, on the right). As shown in Figure 6 for road section 5W, the classical DLM with constant model parameters generates a quite biased one-step forecast due to such a sudden change (black line). However, the proposed method provides adaptive one-step forecasts during and after accidents (red line). The comparison of absolute errors obtained by the classical DLM and the adaptive DLM is shown on the right side of Figure 6. Figure 7 compares the performance of different DLM models for the road section 4E. The result shows the adaptive DLM model obtains more accurate prediction compared to the other DLM models. Figure 8 reports the profile of adaptive optimal signal-to-noise ratios at each time step. We use the standard deviation of travel speed in the normal traffic scenario to estimate the tolerable threshold in (19).

## 5. Conclusions

In this study, we propose an online adaptive DLM algorithm for time series data analysis and forecasting. The proposed method is applied for short-term travel speed forecasts in urban areas based on a microscopic traffic simulator. The

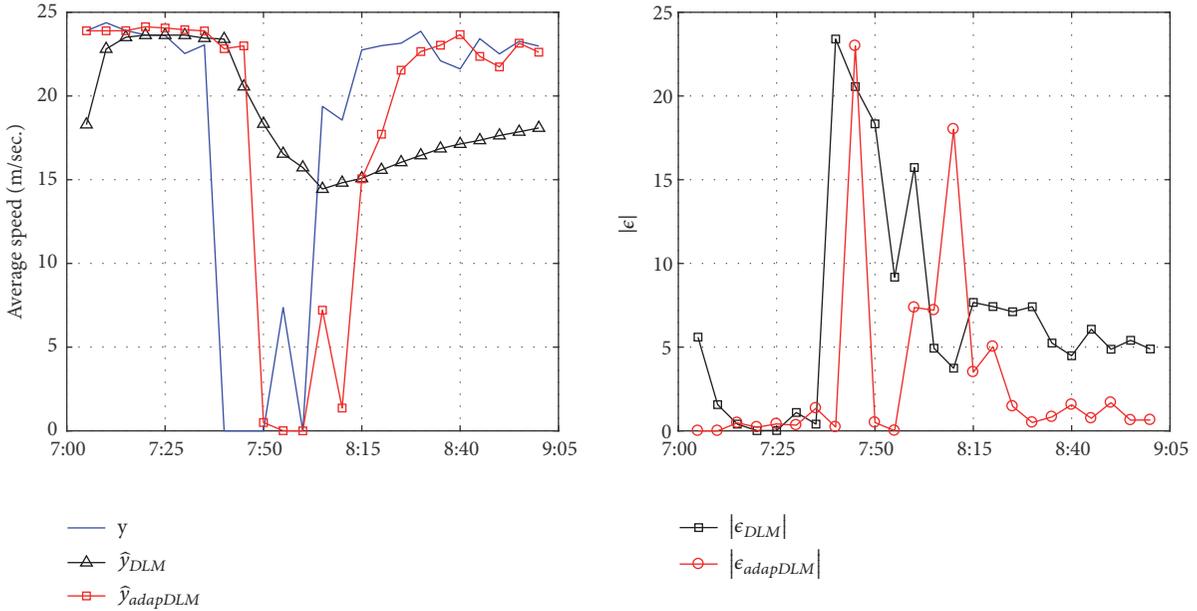


FIGURE 6: Comparison of one-step forecasts and absolute errors for first-order DLM and the adaptive DLM in accident scenarios (road section 5W). *Left*: the one-step forecasts of average speed. *Right*: absolute residuals.

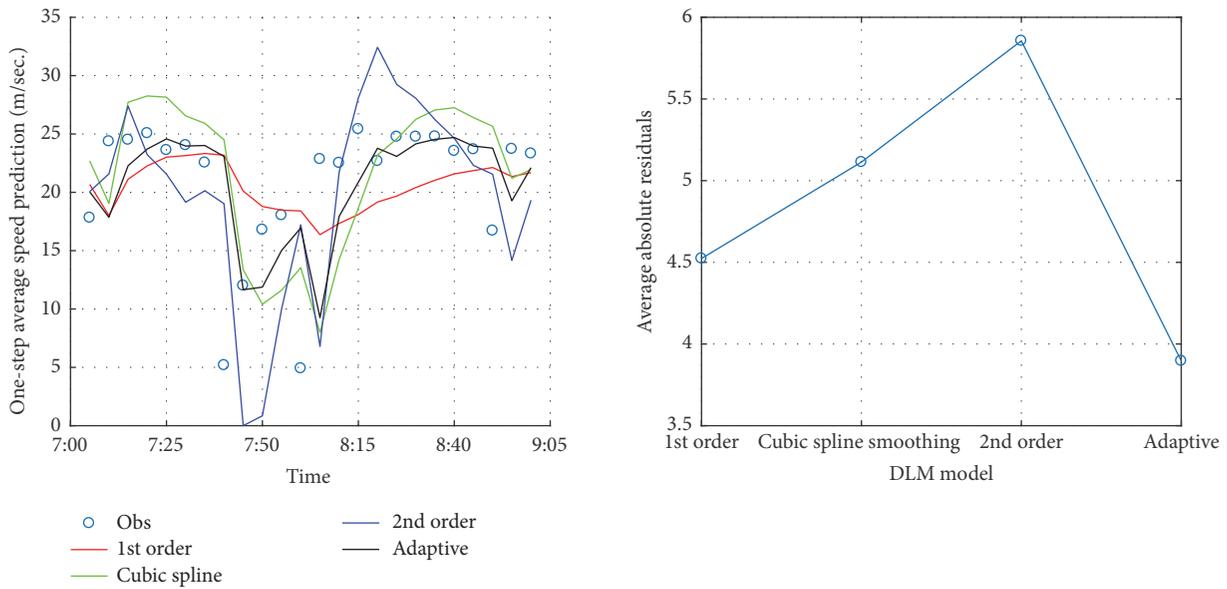


FIGURE 7: Comparison of the performance of different DLM models in unforeseen accident situation (road section 4E).

experiments show the proposed method allows adaptively optimizing its model parameters to improve its prediction accuracy in a continuous way under uncertainty. The proposed method does not need the intervention of experts and can adjust its model error covariance automatically based on feedback information of its one-step prediction errors.

Experimental studies show that our adaptive DLM approach outperforms both autoregressive integrated moving average (ARIMA) and Holt-Winters Exponential Smoothing (ETS) that are both considered to be the main time series

analysis methods employed on this type of problems [28]. We thus consider that this comparison is a reasonable proxy to a comparison with other online models for travel speed prediction that use ARIMA or ETS.

Future extensions concern an adaptive parameter updating scheme design for the state space methods and for more complicated DLMs with seasonal and regression terms. Applications of the proposed method on other time series data would also be beneficial for assessing and improving its performance.

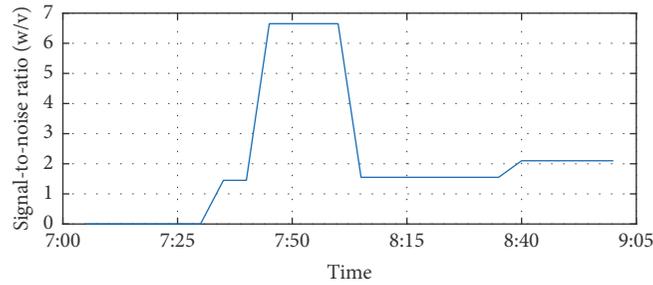


FIGURE 8: Adaptive parameter updating of DLM model in unforeseen accident situation (road section 5W).

## Notations

- $t$ : Index of discretized time intervals,  $t = 1, 2, 3, \dots, T$   
 $\mathbf{x}_t$ : System state at time  $t$   
 $\mathbf{y}_t$ : Observation at time  $t$   
 $\mathbf{v}_t$ : Observation error at time  $t$   
 $V_t$ : Variance of  $\mathbf{v}_t$   
 $F_t$ : Design matrix for observation equation at time  $t$   
 $G_t$ : Evolution matrix of system states at time  $t$   
 $\mathbf{w}_t$ : Forecast error at time  $t$   
 $W_t$ : Variance of  $\mathbf{w}_t$ .

## Data Availability

The data is freely available at <https://github.com/pigne/2019-simulations-DLM>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors are grateful to the support of Luxembourg Institute of Socio-Economic Research (LISER) under the visiting scholar grant.

## References

- [1] T. Ma, "Solving a dynamic user-optimal route guidance problem based on joint strategy fictitious play," in *Game Theoretic Analysis of Congestion, Safety and Security*, pp. 67–89, Springer International Publishing, 2015.
- [2] U. Mori, A. Mendiburu, M. Alvarez, and J. A. Lozano, "A review of travel time estimation and forecasting for advanced traveller information systems," *Transportmetrica A: Transport Science*, vol. 11, pp. 119–157, 2015.
- [3] S. Oh, Y.-J. Byon, K. Jang, and H. Yeo, "Short-term travel-time prediction on highway: a review of the data-driven approach," *Transport Reviews*, vol. 35, pp. 4–32, 2015.
- [4] T. Seo, A. M. Bayen, T. Kusakabe, and Y. Asakura, "Traffic state estimation on highway: A comprehensive survey," *Annual Reviews in Control*, vol. 43, pp. 128–151, 2017.
- [5] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: Where we are and where we are going," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 3–19, 2014.
- [6] C. M. J. Tampere and L. H. Immers, "An extended Kalman filter application for traffic state estimation using CTM with implicit mode switching and dynamic parameters," in *Proceedings of the IEEE Intelligent Transportation Systems Conference*, 2007.
- [7] Y. Wang, M. Papageorgiou, A. Messmer, P. Coppola, A. Tzimitsi, and A. Nuzzolo, "An adaptive freeway traffic state estimator," *Automatica*, vol. 45, no. 1, pp. 10–24, 2009.
- [8] 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.
- [9] Y. Yang, Y. Xu, J. Han, E. Wang, W. Chen, and L. Yue, "Efficient traffic congestion estimation using multiple spatio-temporal properties," *Neurocomputing*, vol. 267, pp. 344–353, 2017.
- [10] S. Fan, M. Herty, and B. Seibold, "Comparative model accuracy of a data-fitted generalized Aw-Rascle-Zhang model," *Networks and Heterogeneous Media*, vol. 9, no. 2, pp. 239–268, 2014.
- [11] K. Y. Chan, T. S. Dillon, J. Singh, and E. Chang, "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and levenberg-marquardt algorithm," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 644–654, 2012.
- [12] L. Chen and C. L. Chen, "Ensemble learning approach for freeway short-term traffic flow prediction," in *Proceedings of the 2007 IEEE International Conference on System of Systems Engineering*, pp. 1–6, San Antonio, Tex, USA, April 2007.
- [13] Y.-S. Jeong, Y.-J. Byon, M. M. Castro-Neto, and S. M. Easa, "Supervised weighting-online learning algorithm for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1700–1707, 2013.
- [14] B. L. Smith, B. M. Williams, and R. 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.
- [15] M. Castro-Neto, Y.-S. Jeong, M.-K. Jeong, and L. D. Han, "Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6164–6173, 2009.
- [16] F. G. Habtemichael and M. Cetin, "Short-term traffic flow rate forecasting based on identifying similar traffic patterns," *Transportation Research Part C: Emerging Technologies*, vol. 66, pp. 61–78, 2016.
- [17] A. Salamanis, G. Margaritis, D. D. Kehagias, G. Matzoulas, and D. Tzovaras, "Identifying patterns under both normal and abnormal traffic conditions for short-term traffic prediction," *Transportation Research Procedia*, vol. 22, pp. 665–674, 2017.
- [18] X. Fei, Y. Zhang, K. Liu, and M. Guo, "Bayesian dynamic linear model with switching for real-time short-term freeway travel

- time prediction with license plate recognition data,” *Journal of Transportation Engineering*, vol. 139, no. 11, pp. 1058–1067, 2013.
- [19] X. Fei, C. C. Lu, and K. Liu, “A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction,” *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1306–1318, 2011.
- [20] Y. Kawasaki, Y. Hara, and M. Kuwahara, “Real-time monitoring of dynamic traffic states by state-space model,” *Transportation Research Procedia*, vol. 21, pp. 42–55, 2017.
- [21] C. Lu and X. Zhou, “Short-term highway traffic state prediction using structural state space models,” *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 18, no. 3, pp. 309–322, 2014.
- [22] A. Stathopoulos and M. G. Karlaftis, “A multivariate state space approach for urban traffic flow modeling and prediction,” *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 2, pp. 121–135, 2003.
- [23] M. West and J. Harrison, *Bayesian Forecasting and Dynamic Models*, Springer, New York, NY, USA, 1997.
- [24] L. Auret and C. Aldrich, “Change point detection in time series data with random forests,” *Control Engineering Practice*, vol. 18, no. 8, pp. 990–1002, 2010.
- [25] G. Comert and A. Bezuglov, “An Online Change-Point-Based Model for Traffic Parameter Prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1360–1369, 2013.
- [26] M. Daumer and M. Falk, “On-line change-point detection (for state space models) using multi-process Kalman filters,” *Linear Algebra and its Applications*, vol. 284, no. 1-3, pp. 125–135, 1998.
- [27] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, “Change-point detection in time-series data by relative density-ratio estimation,” *Neural Networks*, vol. 43, pp. 72–83, 2013.
- [28] L. Moreira-Matias and F. Alesiani, “Drift3Flow: freeway-incident prediction using real-time learning,” in *Proceedings of the IEEE 18th International Conference on Intelligent Transportation Systems*, 571, 566 pages, October 2015.
- [29] E. Ruggieri and M. Antonellis, “An exact approach to Bayesian sequential change point detection,” *Computational Statistics & Data Analysis*, vol. 97, pp. 71–86, 2016.
- [30] E. S. Page, “Continuous inspection schemes,” *Biometrika*, vol. 41, pp. 100–114, 1954.
- [31] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, Wiley-Blackwell, 2015.
- [32] C. C. Holt, “Forecasting seasonals and trends by exponentially weighted moving averages,” *International Journal of Forecasting*, vol. 20, no. 1, pp. 5–10, 2004.
- [33] J. Durbin and S. J. Koopman, *Time Series Analysis by State Space Methods*, vol. 38, Oxford University Press, Oxford, UK, 2nd edition, 2012.
- [34] “Maximum-likelihood method,” in *Encyclopedia of Mathematics*, 2001, [https://www.encyclopediaofmath.org/index.php/Maximum-likelihood\\_method](https://www.encyclopediaofmath.org/index.php/Maximum-likelihood_method).
- [35] G. Petris, S. Petrone, and P. Campagnoli, *Dynamic Linear Models with R*, Springer, New York, NY, USA, 2009.
- [36] J. Kiefer, “Sequential Minimax Search for a Maximum,” *Proceedings of the American Mathematical Society*, vol. 4, no. 3, p. 502, 1953.
- [37] SUMO, *Simulation of Urban Mobility*, 2018, [https://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931\\_read-41000/](https://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931_read-41000/).
- [38] Y. Pigne, G. Danoy, and P. Bouvry, “A platform for realistic online vehicular network management in,” in *IEEE Globecom Workshops*, pp. 595–599, IEEE, 2010.
- [39] Y. Pigné, G. Danoy, and P. Bouvry, “A vehicular mobility model based on real traffic counting data,” in *Communication Technologies for Vehicles*, vol. 6596, pp. 131–142, Springer Berlin Heidelberg, Heidelberg, Germany, 2011.
- [40] Z. Liang and Y. Wakahara, “Real-time urban traffic amount prediction models for dynamic route guidance systems,” *EURASIP Journal on Wireless Communications and Networking*, vol. 85, 2014.
- [41] L. Codeca, R. Frank, and T. Engel, “Luxembourg SUMO traffic (LuST) scenario: 24 hours of mobility for vehicular networking research,” in *Proceedings of the IEEE Vehicular Networking Conference, VNC 2015*, pp. 1–8, Japan, December 2015.



Hindawi

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

