MULTILANE TRAFFIC DENSITY ESTIMATION AND TRACKING

Mikail YILAN[1], Mehmet Kemal ÖZDEMİR[1]

[1]Graduate School of Natural and Applied Sciences, İstanbul Şehir University, İstanbul, Turkey
mikailyilan@std.sehir.edu.tr, kemalozdemir@sehir.edu.tr

Abstract: *As the number of vehicles in roads increases, information of traffic density becomes crucial to municipalities for making better decisions about road management and to the environment for reduced carbon emission. Here, the problem of traffic density estimation is addressed when there is continuous influx of vehicle data. First the traffic density is modeled by the clusters of the speed groups that are centered after Kernel Density Estimation technique is implemented for the probability density function of the speed data. Then, empirical cumulative distribution function of data is found by Kolmogorov-Smirnov Test. A peak detection algorithm is used to estimate speed centers of the clusters. Since the estimation model has linear and non-linear components, the estimation of variance values and kernel weights are found by a nonlinear Least Square approach with separation of parameters property. Finally, the tracking of former and latter estimations of a road is calculated by using Scalar Kalman Filtering with scalar state - scalar observation generality level. For all example data sets, the minimum mean square error of kernel weights is found to be less than* 0.002 *while error of mean values is found to be less than* 0.261.
Keywords: *Traffic Density, Kernel Density Estimation, Kolmogorov-Smirnov Test, Nonlinear Least Square, Scalar Kalman Filter.*

1. Introduction

The Estimation and prediction of the traffic density is necessary to prevent citizens from congested traffic. If the decision makers have the knowledge of the current and future traffic reports, municipalities would come up with better solutions against the traffic problem. Also, drivers could have better route options to follow while driving. In doing so, they will spend less gasoline and time, and hence low carbon emission and less air pollution will result in. Although there are many different ways to estimate the traffic density, Kernel Density Estimation (KDE) is one of the best estimation techniques since cars that go on the same road with different speeds on different lanes can be better represented by KDE [1]. On the other hand, parametric and non-parametric approaches form the two types of estimation techniques. Former one has a fixed number of parameters and the latter one has an increasing number of parameters when the training data size becomes larger [2]. Since the estimation of traffic parameters needs to cope with continuously incoming data from the field, KDE, which is non-parametric, is exploited to better describe the problem [3]. Gaussian distribution performs well to represent real-time data, and therefore the sample data is typically modeled as normally distributed. In [4] and [5], which are the initial work of this study, KDE was used to derive the probability density function (PDF) of the received data. As mentioned above, KDE reveals various traffic scenarios accurately and it is helpful for theoretical improvements when it is compared with other methods for the PDF expression. In the aforementioned studies, the cumulative distribution function (CDF) was found by Kolmogorov-Smirnov (KS) test, which is less affected by the existence of outliers when compared with other tests [6].

When traffic density is estimated, its data can be thought of a collection of clusters. The clusters are formed by three parameters: kernel weights that show the corresponding cluster's weight among all available clusters, speed centers, and bandwidths. In our first study [5], kernel weights are estimated by using KDE, KS test, and linear Least Square approaches, while the other two parameters were treated as constants. In our earlier study [4], all of these parameters were taken as non-constants. In the first step of estimations, a peak detection algorithm (PDA) over the smoothed version of the PDF was utilized for the estimation of mean values. Nonlinear LS (NLS) with separation of parameters approach was applied successively to estimate variance values and kernel weights. First, speed center's variance and then its kernel weights were estimated. After these estimations, the next speed center's bandwidth and its kernel weight were estimated, and so on. Linear search method that gives accurate results and Newton-Raphson (N-R) Method that reaches to the solution in a quite shorter time were exploited in the NLS approach [7].

In this paper, an extension to the work in [4] will be presented. For the same road, if new data arrives in addition to the already existing data that is used for the estimation, we adopt the tracking of the estimated parameters instead of

Received on: 30.03.2016
Accepted on: 17.05.2016

using the new large data set and recalculating all the parameters needed to make a fresh estimate. In accomplishing this, first an initial estimate is obtained from the new and small sized data. Then, to get an overall estimate, the old and the new estimates are combined. Hence, a forgetting factor is utilized for the tracking of the old and the new estimates so that more weight is given to the newly arrived data rather than updating them according to their number of samples.

In the next part, in Section II, the model of the system will be verbalized. In Section III, numerical estimations and their tracking will be presented and the system performance will be assessed. In Section IV, in light of these efforts, the deductions will be made. Finally in Section V, a summary of conclusions and future prospective will be given.

2. The Model

The first two subsections of this section includes some equations that are formulated in [4] but we briefly present them here since they are also used for the tracking algorithm adopted in this paper.

2.1. Finding PDF with KDE and Empirical CDF with KS Test

For a given N independent samples, let $x \equiv \{X_1, \ldots, X_N\}$ comes from a continuous PDF f, which is defined on X. Gauss KDE can then be defined as follows [8]:

When the mean of each data sample is X_i and the corresponding variance is σ, then the Gauss Kernel PDF is

$$\hat{f}(x;\sigma) = \frac{1}{N}\sum_{i=1}^{N}\varphi(x, X_i;\sigma), \qquad x \in R \tag{1}$$

where

$$\varphi(x, X_i;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-X_i)^2/(2\sigma)} \tag{2}$$

When the contribution of the Gauss Kernels are different, then (1), can be re-expressed by including kernel weights, α_i as:

$$\hat{f}(x;\sigma,\alpha) = \frac{1}{N}\sum_{i=1}^{N}\alpha_i\varphi(x, X_i;\sigma), \qquad x \in R \tag{3}$$

where

$$0 \le \alpha_i \le 1 \quad \text{and} \quad \sum \alpha_i = 1 \tag{4}$$

Although there are lots of approaches for the representation of a PDF, when KDE is preferred for mapping, representation via KDE would be more suitable for visualization and a better theoretical background [8]. Hence, the smoothed version of data is treated to represent PDF with KDE since one of the essential purposes of KDE is to produce a smooth density surface over a 2-D geographical space [9].

The next step is the expression of empirical CDF and here, KS Test is chosen for the implementation since it is less disturbed by the outliers. KS Test initially detects the difference among the real and the empirical speed values of data set and also how much they are close to each other. Since the distribution of speed values is assumed to be Gaussian, this assumption can be examined in the CDF plots. The methodology is to arrange all datum in the data with an increased order and then to rescale them [10]. As expected, the CDF plot reaches to unity when the last datum is processed.

When F is defined as CDF, $\hat{F}$ would be the empirical CDF and $F(x)$ is counted as equal to $F_0(x)$ as a hypothesis for all $x \in R$ values. Then, KS Test statistics are defined as supremum of the difference between $\hat{F}(x)$ and $F_0(x)$ as described in [10].

2.2. Determination of Speed Centers via PDA and Estimation of Variance and Kernel Weights with Nonlinear LS Method

The clusters in a data set are decided according to speed centers that correspond to different regions in the PDF. Therefore, primarily all mean values should be estimated via PDA. The algorithm is applied to the PDF, and the speed centers are found straightforwardly. First, the derivative of the CDF is calculated that produces the PDF. Later, the resulting values are smoothed to get the PDF for the algorithm implementation. With these, the PDA generates accurate mean values since it examines every datum in the data set and catches the peak values. Corresponding mean values of such peak values are the centers of the clusters. After determination of mean values, their variances and kernel weights are estimated by using separability of parameters property of NLS. In the model, α is linear and σ is nonlinear with respect to the model as seen in (3). The approach is to estimate α via LS in terms of σ and then to estimate σ. In order to achieve this, the following equation (LS error) should be minimized for α [7];

$$J(\sigma,\alpha) = (x - H(\sigma)\alpha)^T(x - H(\sigma)\alpha) \tag{5}$$

Here x corresponds to F values of empirical CDF, and hence F will be used instead of x. The estimation for α is then:

$$\hat{\alpha} = \left(H^T(\alpha)H(\alpha)\right)^{-1}H^T(\alpha)F \tag{6}$$

Then, by replacing $\hat{\alpha}$ into above LS error (5), we get:

$$J(\sigma,\hat{\alpha}) = F^T\left(I - H(\alpha)\left(H^T(\alpha)H(\alpha)\right)^{-1}H^T(\alpha)\right)F \tag{7}$$

Thus, minimization of $J(\sigma,\hat{\alpha})$ is the same as maximization of the following equation over α:

$$\max_{\alpha}\left[F^T H(\alpha)\left(H^T(\alpha)H(\alpha)\right)^{-1}H^T(\alpha)F\right] \tag{8}$$

We apply the above maximization for each speed cluster with the following intermediary variables:

$$H_i = \frac{1}{2}\left(1 + erf\,\frac{x-\mu_i}{\sigma\sqrt{2}}\right) \quad (9)$$

$$F = \sum \alpha_i H_i \quad (10)$$

$$F_i = \left(\alpha_i - \sum_{j=1}^{i-1} \alpha_j\right) H_i \quad (11)$$

where F_i and H_i are for F and H, respectively, for the i^{th} cluster and μ_i is the mean speed value for the same cluster. Then the kernel weights α_is for each cluster can be determined as

$$\alpha_i = (H_i^T H_i)^{-1} H_i^T F_i - \sum_{j=1}^{i-1} \alpha_j \quad (12)$$

Initially, the bandwidth of the mean values is calculated via (8) by using linear search method or Newton-Raphson Method [4]. By then inserting this into (12), the kernel weight of the corresponding speed center is determined. Since the mean values are already found before the application of NLS method, the same procedure for the estimation of the variance and kernel weight is repeated for each speed center, thereby resulting in a successive estimation process.

2.3. Tracking of Traffic Density Estimation

Tracking is very much needed when newly arrived data needs to be processed in addition to the past data. Instead of going back to the initial state of estimation of the parameters by using all the existing and the newly arrived data, the estimation of the parameters is just updated with the arrival of new data. Hence, the final estimates are like reaching a consensus between already estimated parameters and newly estimated ones. Moreover, the importance of the old and new data is not the same for the estimation, because new data has more emphasis on estimation and is seen as more probable to convey the current traffic scenario. Therefore, on the contrary to just reordering estimation results according to their number of samples in data sets, the use of a forgetting factor is necessary to improve the tracking capability in time varying parameter estimation [11]. Forgetting factor can be defined as the concept of forgetting in which older data is gradually scrapped by taking into consideration of more recent information [12]. The main idea behind this concept is to give less weight to older data and more weight to the new one [12].

In this study, Scalar Kalman Filter (KF) has been used for tracking. Its scalar state - scalar observation $(s[n-1], x[n])$ generality level is chosen as an approach. The scalar state and the scalar observation equations are as follows [7]:

$$s[n] = \lambda s[n] + u[n] \quad n \geq 0 \quad (13)$$

$$x[n] = s[n] + w[n] \quad (14)$$

where λ is called the forgetting factor with $0 < \lambda < 1$, $u[n]$ is White Gaussian Noise (WGN) with $u[n] \sim \mathcal{N}(0, \sigma_u^2)$, $w[n] \sim \mathcal{N}(0, \sigma_w^2)$, and $s[-1] \sim \mathcal{N}(\mu_s, \sigma_s^2)$. $w[n]$ differs from WGN only in that its variance is allowed to change in time. Further assumption is the independence of $u[n]$, $w[n]$, and $s[-1]$.

$s[n]$ is estimated based on the data set $\{x[0], x[1], \dots, x[n]\}$ as n increases, and this process is simply a type of filtering. KF approach calculates the estimator $\hat{s}[n]$ subjected to the estimator for the previous sample $\hat{s}[n-1]$ and thus, it is recursive in nature [7].

With $n \geq 0$, the scalar KF equations (Prediction (Pr), Minimum Prediction MSE (Min Pre MSE), Kalman Gain (KG), Correction (Cr), Minimum MSE (Min MSE), respectively) for tracking are as follows:

Pr: $\hat{s}[n|n-1] = \lambda \hat{s}[n-1|n-1]$ (15)

Min Pr MSE: $M[n|n-1] = \lambda^2 M[n-1|n-1] + \sigma_u^2$ (16)

KG: $K[n] = (M[n|n-1])/(\sigma_w^2 + M[n|n-1])$ (17)

Cr: $\hat{s}[n|n] = \lambda \hat{s}[n|n-1] + K[n](x[n] - \hat{s}[n|n-1])$ (18)

Min MSE: $M[n|n] = (1 - K[n])M[n|n-1]$ (19)

3. Numerical Calculations

In this part, the system will be tested with 3 examples: the first one examines tracking with the change only in speed centers, the second one evaluates the tracking of kernel weights' changes, and finally the last one investigates what happens if the all variables have new different values. To simulate the given scenarios, Data Set 1 is produced and estimated firstly and then used in all three examples. For the first example, Data Set 2 and Data Set 3 are also created. For the second example, tracking of kernel weights is performed by using Data Set 1 and Data Set 4. In the last example, Data Set 1 and Data Set 5 are used for the estimations and tracking.

Before we further proceed, we will state how some parameters in the tracking process are chosen. For example, the forgetting factor λ is calculated as a ratio of number of samples in the first data set and the number of all samples. By using the forgetting factor, the sample numbers are used implicitly in the tracking equation, however, the tracking is simply neither updating the overall estimation according to the number of samples nor giving equal weights to both old and new estimates. As the examples will show, the final estimates are closer to the estimates based on the newly arrived data rather than the estimates from the old data. Assuming that the traffic data is obtained via GPS data, and since the accuracy of GPS data is at least 95% according to GPS Standards [13], error values σ_u^2 and σ_w^2 are assumed to 0.05. For the tracking of mean, instead of $\hat{s}[n-1|n-1]$, Data Set 1's mean estimation and instead of $x[n]$, new data set's mean estimation are used. For the kernel weight's tracking, kernel weight estimations of the aforementioned data sets are taken into calculations. The MMSE (minimum mean square error) of each variable is calculated separately since the system is scalar.

A traffic scenario with 3 speed centers is assumed. The performance of the approach will be assessed for the estimated means, variances, and kernel weights as well as tracking. The assumed scenario for the first data set has the following parameters with the number of samples $N = 10000$:

$$\mu_1 = 50 \quad \mu_2 = 70 \quad \mu_3 = 100$$

$$\sigma_1^2 = 6 \quad \sigma_2^2 = 7 \quad \sigma_3^2 = 5$$

$$\alpha_1 = 0.3 \quad \alpha_2 = 0.5 \quad \alpha_3 = 0.2$$

It is needed to be emphasized that differently from [4], in addition to Gaussian distribution variance values, the contribution of GPS allowable error is added as variance. For the given example, uniformly distributed additional variance values are 2.5, 3.5, and 5, respectively. The estimation of Data Set 1's mean values via a peak detection algorithm are as follows:

$$\hat{\mu}_1 = 50.1489 \quad \hat{\mu}_2 = 69.9231 \quad \hat{\mu}_3 = 100.1197$$

The estimation results are very close to the real values as the MMSE is 0.04240:0424. Also, the MMSE of each speed center's estimation of Data Set 1 are 0.2220, 0.0059, and 0.0143, respectively. The variances and kernel weights are estimated by using two methods as explained in Section II-B. For N-R Method, which reaches accurate results quicker, the estimated values are as follows:

$$\hat{\sigma}_1^2 = 9.5088 \quad \hat{\sigma}_2^2 = 10.2556 \quad \hat{\sigma}_3^2 = 22.8845$$

$$\hat{\alpha}_1 = 0.3038 \quad \hat{\alpha}_2 = 0.4674 \quad \hat{\alpha}_3 = 0.2278$$

When the error values are analyzed, kernel weights and speed centers have less error when compared to variance's values. However, the estimation of variances is an intermediate step before the estimation of the kernel weights. Although variance estimation provides useful information about the traffic density, the speed centers and kernel weights are more critical in assessing multi-lane traffic density. The MMSE of kernel weights is 0.0019 and also the MMSE of each kernel weights of Data Set 1 are 1.4440×10^{-5}, 0.0011, and 0.0008, respectively. For the linear search method, which takes longer time but that generally provides more accurate results, the estimated variances and kernel weights are as follows:

$$\hat{\sigma}_1^2 = 9.5090 \quad \hat{\sigma}_2^2 = 10.2560 \quad \hat{\sigma}_3^2 = 22.8840$$

$$\hat{\alpha}_1 = 0.3002 \quad \hat{\alpha}_2 = 0.4619 \quad \hat{\alpha}_3 = 0.2379$$

with MMSE of 0.0029.

As can be seen from the estimated values, the proposed approach can accurately estimate the targeted parameters as MMSE values are acceptably small for the traffic density estimation.

For the first example, new data sets (Data Set 2 and Data Set 3) are produced with a 5 km/h increase in speed centers while keeping the other parameters unchanged. These data sets have their number of samples as $N = 1000$, and by doing so we will examine the results of tracking by repeating the same procedure. Here, the expectation is that the second tracking would be closer to the speed centers of new data set than the first tracking. For Data Set 2, estimation of mean values and the overall system's corrected speed centers are as follows:

$$\hat{\mu}_1 = 54.9107 \quad \hat{\mu}_2 = 74.8072 \quad \hat{\mu}_3 = 105.2616$$

$$\hat{\mu}_{cor1} = 53.9121 \quad \hat{\mu}_{cor2} = 71.9633 \quad \hat{\mu}_{cor3} = 103.1484$$

The MMSE of mean estimation of Data Set 2 is 0.1561. Again, the estimation is very close to the real values. Also, the MMSE of each speed center's estimation of Data Set 2 are 0.0302, 0.0431, and 0.0828, respectively. As seen from

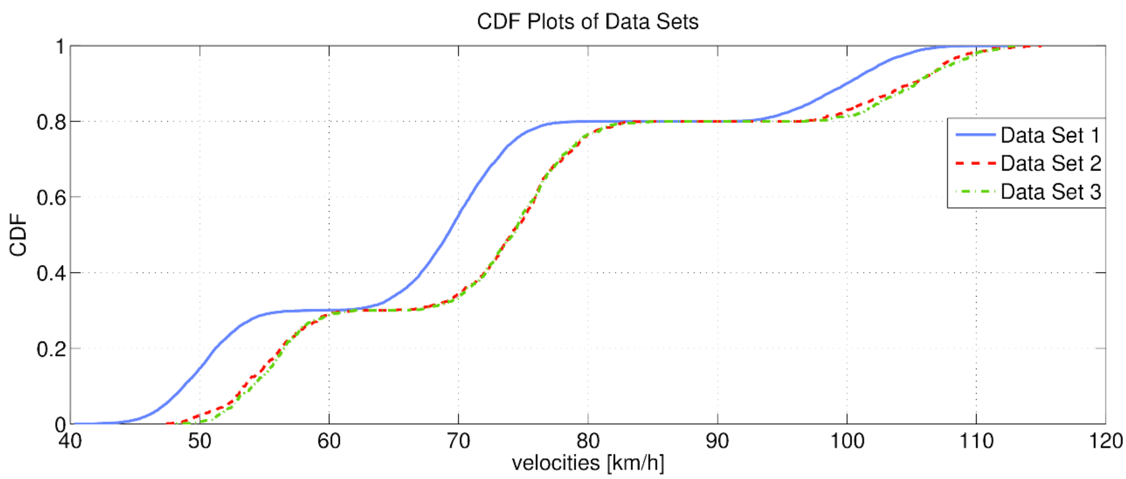

Figure 1. CDF plots of three data sets

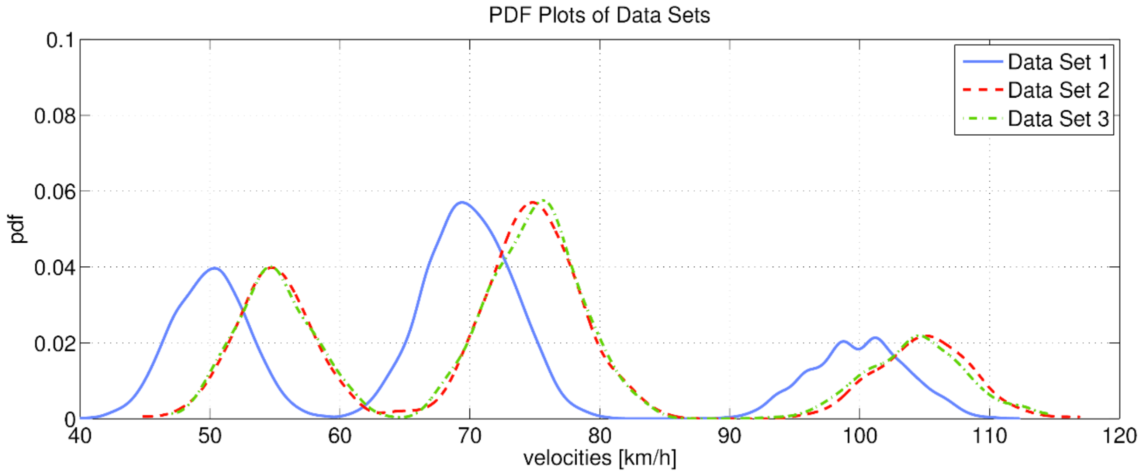

Figure 2. PDF plots of three data sets

results of tracking, corrected mean values are not linearly calculated by simply taking the number of samples in data sets. It is also observed that tracking represents real traffic scenarios better since the new data sets have more effect in the evaluation of the current estimates even though they bear less number of samples. It is obvious that if the initial estimates of the firstly received data are more accurate, then the results are closer to the real values. The estimation of first speed center has more error than the estimation of the second speed center. Thus, the second one has corrected mean values closer to its former estimation (50.1489) than the corrected version of the first mean value's closeness to its former estimation (69.9231). Data Set 3 also has the sameparameter values with Data Set 2. Estimation of its mean values and corrected mean values of the overall system that consists of all three data sets including the results of first tracking are as follows:

$$\hat{\mu}_1 = 54.6585 \quad \hat{\mu}_2 = 75.3391 \quad \hat{\mu}_3 = 105.0514$$

$$\hat{\mu}_{cor1} = 54.2267 \quad \hat{\mu}_{cor2} = 74.7701 \quad \hat{\mu}_{cor3} = 104.6992$$

The MMSE of mean estimation of Data Set 3 is 0.2342. As expected, new corrected values of speed centers are higher than the former ones. The illustration of the change in mean values and their kernel weights and variances of data sets can be observed in Figs. 1 and 2.

In the second example, only kernel weights will change and we will track their values. Data Set 4's kernel weights for $N = 1000$ number of samples are given as follows:

$$\alpha_1 = 0.5 \quad \alpha_2 = 0.4 \quad \alpha_3 = 0.1$$

Since in our case N-R Method's MMSE is less than linear search one, estimation of kernel weights and their tracking are as follows:

$$\hat{\alpha}_1 = 0.5154 \quad \hat{\alpha}_2 = 0.3863 \quad \hat{\alpha}_3 = 0.0983$$

$$\hat{\alpha}_{cor1} = 0.3961 \quad \hat{\alpha}_{cor2} = 0.4027 \quad \hat{\alpha}_{cor3} = 0.1466$$

The MMSE of kernel weights estimation of Data Set 4 is 4.2658×10^{-4}. As seen from corrected kernel weights results, the error of each values of first data set's estimation is correlated with final calculated values. The difference between Data Sets 1 and 4 can be seen in Figs. 3 and 4.

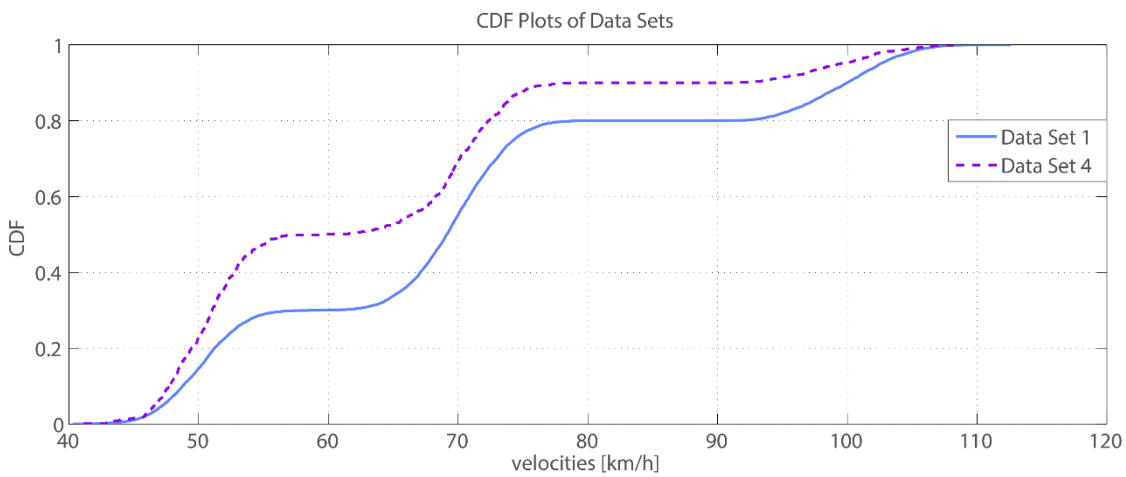

Figure 3. CDF plots of Data Set 1 and Data Set 4

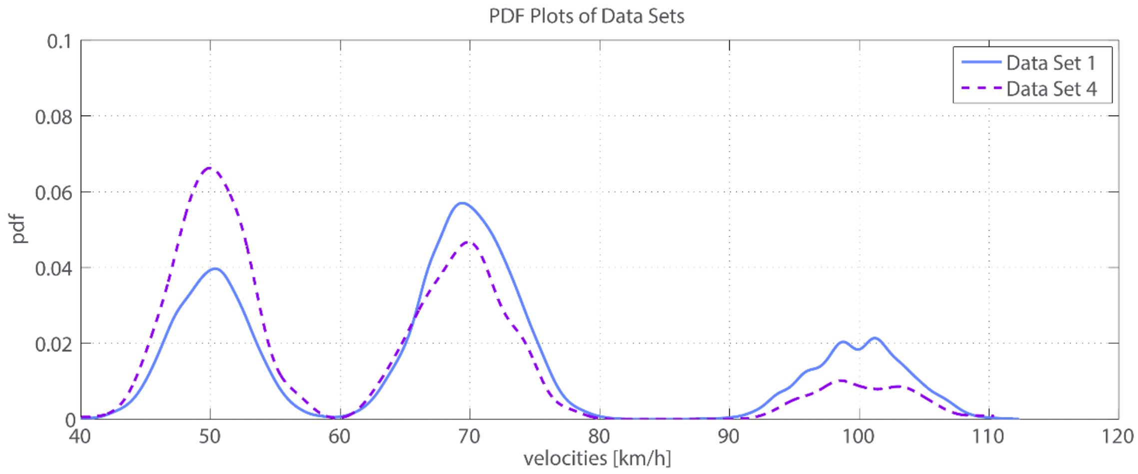

Figure 4. PDF plots of Data Set 1 and Data Set 4

In the last example, all parameters' values of Data Set 1 are changed, and their tracking is calculated for $N = 1000$ sample number. The estimation of mean values via PDA and the estimation of kernel weights and variance values with N-R Method including assumed sample system parameters are given in Table I. Also tracking results of Data Set 5 are shown in the table. The MMSE of mean estimation is 0.2609 and the MMSE of kernel weight estimation is 1.6213×10^{-4}. As seen from all examples, the system has performed well for not only the change of a single parameter but also for the change of all parameters. The PDF and CDF plots of Data Sets 1 and 4 are shown in Figs. 5 and 6.

4. Assessment

The system is tested with three different examples and for all three, it performed well by giving out the desired results. Chosen speed values are relatively middle and high speed levels for driving standards. The model was also examined in low and high speed values, and it performed well for estimations and their tracking. In addition to the work in [4], this work dealt with not only estimations of kernel weights, mean values and variance values but also their tracking. While in estimating variances, N-R Method reaches the results very quickly and its performance is comparable to the linear search method results. For example, for every estimation process of each variance value, linear search method needs more than 100 thousand multiplications to get maximum values in (8), while N-R Method reaches maximum value in less than 10 iteration even though its evaluation needs some heavy computation [4]. Scalar KF with scalar state - scalar observation generality level has achieved preferred outcomes instead of just linear calculations of two estimation results.

Table 1. Example 3 Parameters, Results of Estimation and Tracking

Parameter # \ Name	μ	σ^2	α	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\alpha}$	$\hat{\mu}_{cor}$	$\hat{\alpha}_{cor}$
1	55	5+2.75	0.1	54.7625	6.0102	0.0918	53.7797	0.1838
2	75	6+3.75	0.3	74.9063	12.3964	0.2985	72.0373	0.3523
3	105	7+5.25	0.6	105.4424	19.1120	0.6096	103.3024	0.4331

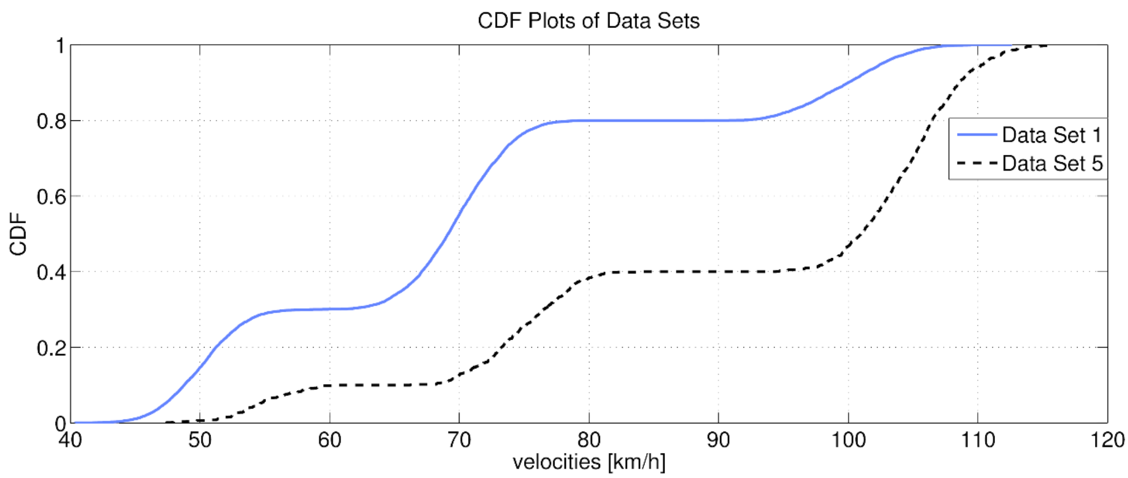

Figure 5. PDF plots of Data Set 1 and Data Set 5

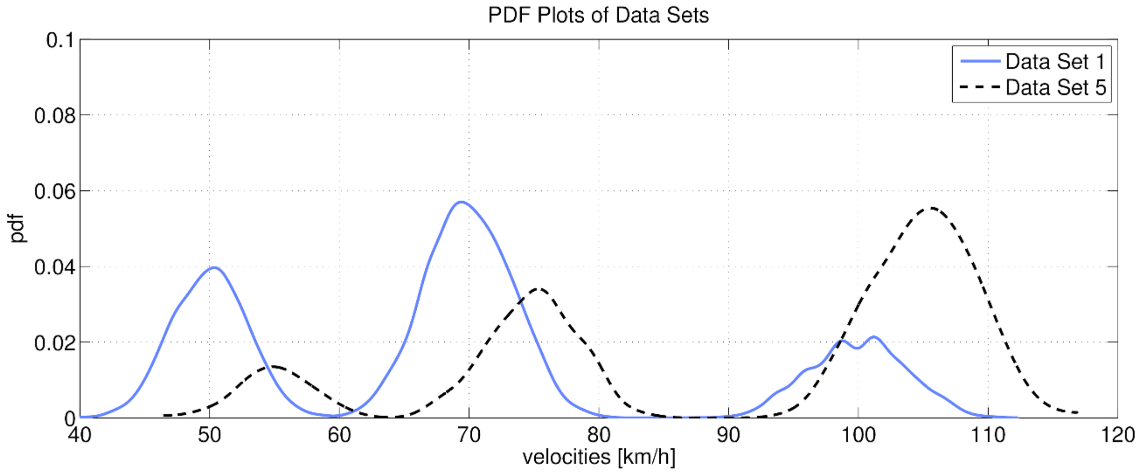

Figure 6. CDF plots of Data Set 1 and Data Set 5

By using both Netwon-Raphson and Linear Search methods, average running time for parameter estimations of old data and new data, and their tracking is approximately 2 minutes and 10 seconds. However, here, estimation of the same parameters are made twice to compare these two methods. As mentioned in Section I, although Newton-Raphson method has computational complexity, it is faster than Linear Search method. If we use only Newton-Rapson method to estimate old data and new data, then perform the tracking, the average running time reduces to approximately 55 seconds. The computer used has a 1.70 GHz CPU, 4.00 GB RAM, and 128.00 GB memory.

The simulation results indicate that the error for mean estimation is less 0.261, while it is less than 0.002 for kernel weights when N-R Method is used. Here the variances are calculated as intermediate variables in estimating the kernel weights. These error rates are considered to represent multilane traffic condition accurately when compared with other studies in the literature such as [14] and [15]. We can use the following mean-percentage error formula that is used in [14] to compare error rates as

$$E_{MPE} = \frac{1}{M}\sum_{i=1}^{M}\left|\frac{\alpha_i \hat{\alpha}_i}{\alpha_i}\right| \quad (14)$$

Here, M is the number of total speed centers and it is equal to 3 in the current study.

The error in [14] is around 13% and the error in [15] is around 10%. Meanwhile, the error rate of mean values in this study reaches a maximum value of 0.11% when (20) is used. If we modify the equation (20) for error rate of kernel weight estimation, its maximum becomes 2.04%. As seen in the Section III, for different data sets and examples (5 data sets for 3 different examples), the error values do not change much and it is sufficiently less than the prior art. Thus, by using the proposed approach, an accurate multi-lane traffic density estimation and its tracking are realized.

5. Conclusion and Future Study

Traffic density estimation and its prediction play a crucial role in managing the traffic on the roads. The overall outcome prevents the drivers from traffic congestion and wasted-time and therefore is very beneficial to both the drivers and the management bodies of municipalities. In this study, multi-lane traffic density estimation has been conducted by estimating the speed centers, bandwidths, and kernel weights of clusters, which represent a group of moving vehicles in a given road and lanes. For this, the PDF of the input data is found by implementing Kernel Density Estimation. Then Kolmogorov-Smirnov Test is used to find empirical CDF. Thereafter, mean values are estimated via a peak detection algorithm and then variance values and kernel weights are estimated successively by using separation of parameters property of nonlinear Least Square Method that is applied with linear search method and Newton-Raphson approaches. As an extension to [4], for the same road, tracking of former and new estimations with less amount of data is determined by using Scalar Kalman Filter with scalar state - scalar observation generality level. The roads' traffic density estimation is then updated with the newly calculated values. Three different sample cases representing a) change in speed centers, b) change in kernel weights, and c) change in all parameters, i.e. mean values, kernel weights, and variance values are analyzed in order to validate the proposed model. It is observed that the proposed estimator and the tracking algorithm perform very well when compared with the state of the art. This current study can further be extended to the prediction of the multi-lane traffic density for a given time interval, say daily or weekly.

6. References

[1] Laxhammar R., Falkman G., and Sviestins E., "Anomaly Detection in Sea Traffic - A Comparison of the Gaussian Mixture Model and the Kernel Density Estimator," in *Information Fusion, 2009, 12th International Conference* on, July 2009, pp. 756–763.

[2] Murphy K. P., *Machine Learning: A Probabilistic Perspective*, MIT press, 2012.

[3] Tabibiazar A. and Basir O., "Kernel-Based Optimization for Traffic Density Estimation in ITS," in *Vehicular Technology Conference (VTC Fall)*, 2011 IEEE, Sept 2011, pp. 1–5.
[4] Yılan M. and Özdemir M. K., "Traffic Density Estimation via KDE and Nonlinear LS [submitted, pending]," in *Turkish Journal of Electrical Engineering & Computer Sciences*, 2016.
[5] Yılan M. and Özdemir M. K., "A Simple Approach to Traffic Density Estimation by using Kernel Density Estimation," in *Signal Processing and Communications Applications Conference (SIU), 2015 23th*, May 2015, pp. 1865–1868.
[6] Jäntschi L., Bolboaca S. D. et al., "Distribution Fitting 2. Pearson-Fisher, Kolmogorov-Smirnov, Anderson-Darling, Wilks-Shapiro, Kramer-von-Misses and Jarque-Bera statistics," *Bulletin UASVM Horticulture*, vol. 66, no. 2, pp. 691–697, 2009.
[7] Kay S. M., *Fundamentals of Statistical Signal Processing: Estimation Theory*, Pearson Education, 2013, vol. 1.
[8] Botev Z. I., Grotowski J. F., Kroese D. P. et al., "Kernel Density Estimation via Diffusion," *The Annals of Statistics*, vol. 38, no. 5, pp. 2916–2957, 2010.
[9] Xie Z. and Yan J., "Kernel Density Estimation of Traffic Accidents in a Network Space," *Computers, Environment and Urban Systems*, vol. 32, no. 5, pp. 396–406, 2008.
[10] Djuric P. and Miguez J., "Model Assessment with Kolmogorov-Smirnov Statistics," in *Acoustics, Speech and Signal Processing, 2009, ICASSP 2009, IEEE International Conference* on, April 2009, pp. 2973–2976.
[11] Song S., Lim J.-S., Baek S., and Sung K.-M., "Gauss Newton Variable Forgetting Factor Recursive Least Squares for Time Varying Parameter Tracking," *Electronics Letters*, vol. 36, no. 11, pp. 988–990, May 2000.
[12] Vahidi A., Stefanopoulou A., and Peng H., "Recursive Least Squares with Forgetting for Online Estimation of Vehicle Mass and Road Grade: Theory and Experiments," *Vehicle System Dynamics*, vol. 43, no. 1, pp. 31–55, 2005.
[13] DoD U., "Global Positioning System Standard Positioning Service Performance Standard," *Assistant Secretary of Defense for Command, Control, Communications, and Intelligence*, 2008.
[14] Munoz L., Sun X., Horowitz R., and Alvarez L., "Traffic Density Estimation with the Cell Transmission Model," in *American Control Conference, 2003. Proceedings of the 2003*, vol. 5, June 2003, pp. 3750–3755 vol.5.
[15] Sun X., Munoz L., and Horowitz R., "Highway Traffic State Estimation using Improved Mixture Kalman Filters for Effective Ramp Metering Control," in *Decision and Control, 2003. Proceedings, 42nd IEEE Conference* on, vol. 6, Dec 2003, pp. 6333–6338 Vol. 6.

Note:

Mikail Yılan completed his BSc in electrical-electronics engineering at Boğaziçi University, İstanbul Turkey in '14. He received his MSc from İstanbul Şehir Univesity, Turkey in '16 from electronics and computer engineering. Starting from February '16, he works in industry for IstLink Communications in the area of LTE-A systems. His research interests are traffic density estimation and interference cancellation in HetNets in LTE-A.

Mehmet Kemal Özdemir completed his BSc and MSc in electrical engineering at METU, Ankara Turkey in '96 and in '98, respectively. He received his PhD from Syracuse University, Syracuse, USA in '05 from electrical engineering. Starting from '99, he worked in industry in the area of broadband communication with the focus on CATV and 4G wireless systems. Dr. Ozdemir is currently with Istanbul Sehir University, Istanbul, Turkey as an Asst. Prof at the EE department. His research interests are receiver algorithms for OFDM based systems, traffic density estimation, 5G receiver design, and FM band directional channel modeling.