CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Real-time Density Estimation on Freeway with Loop Detector and Probe Data

Tony Z. Qiu, Xiao-Yun Lu,
Andy H. F. Chow, Steven Shladover

California PATH Working Paper
UCB-ITS-PWP-2009-6

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation, and the United States Department Transportation, Federal Highway Administration.

Report for Task Order 6224

April 2009

ISSN 1055-1417

CALIFORNIA PARTNERS FOR ADVANCED TRANSIT AND HIGHWAYS

Real-time Density Estimation on Freeway with Loop Detector and Probe Vehicle Data

PATH Working Paper

Tony Z. Qiu, Xiao-Yun Lu, Andy H. F. Chow, Steven Shladover

FHWA-EXARP and Caltrans TO6224

April 20, 2009

Key Words

Density Estimation, Vehicle Infrastructure Integration (VII), Loop Detector, Probe Vehicle

Abstracts

Density, speed and flow are the three critical parameters for traffic analysis. Traffic management and control with high performance require accurate estimation/prediction of distance mean speed and density for large spatial and temporal coverage. Speed, including time mean speed and distance mean speed, and flow estimation are relatively easy to be measured and estimated in the practical site, but accurate density estimation is very difficult. Inductive loop detector systems have been widely deployed, it makes better sense to fully adopt available infrastructure to achieve required traffic measurement. As a new promising technology for transportation system, Vehicle Infrastructure Integration (VII) is developing rapidly with the market penetration of cell phone and GPS systems. This report proposed a method for real-time estimation of density using synchronized loop detector data and VII probe vehicle data. Berkeley Highway Laboratory (BHL) loop detector data and the field collected Probe Vehicle data have been used in the method validation. Density estimated from the vehicle-by-vehicle trajectory tracking in Next Generation Simulation (NGSIM) data has also been used as the second data source for validating the algorithm. Comparison of the two results – that form the loop and VII probe vehicle data and that from NGSIM data, showed that they are very close except a small offset which needs further investigation.

Acknowledgements

This work was supported by the Federal Highway Administration (FHWA) Exploratory Advanced Research Project (Cooperative Agreement DTFH61-07-H-00038) with Match Funding from the California Department of Transportation (Caltrans) TO6224. The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of FHWA and the State of California. This report does not constitute a standard, specification, or regulation.

Table of Contents

Page

List of Figures and Tables

1 Introduction

Density, speed and flow are the three critical parameters for traffic analysis. Traffic management and control with good performance requires the real-time estimation of distance mean speed and density as input for large spatial and temporal coverage of roadway network. It would be better that density can be measured and estimated lane by lane.

Generally, it is believed that traffic density estimation or prediction is more difficult than the estimation or prediction of distance mean speed and flow. This is partly because the most popularly used sensors are inductive loop detectors which are point sensors while density is a range concept. The disturbances to density estimation include location, weather, and vehicle types, etc. Further, lane changing behavior affects lane-wise density significantly. As a matter of fact, both lane-wise distance mean speed and density are very difficult to be measured/estimated accurately in real-time although they are required in Coordinated Ramp Metering (CRM) and Variable Speed Limit (VSL) control.

This report is organized as the following: Section 2 is for literature review. Section 3 lists the traditional method only using loop detector data, and presents the proposed methodology based on available loop detector data and VII probe vehicle data. Section 4 details the results from our field test, and empirical results and their implications on developing traffic control strategies are also discussed. Finally, Section 5 gives some concluding remarks and future study.

1.1 Density Estimation

As an instantaneous and range concept, traffic density is primarily defined by looking into a snapshot photo of the traffic by an aerial camera along a stretch of freeway (Greenshields, 1933; Greenshields, 1935). The density is the number of vehicles divided by the length of the stretch. Average density over lanes is naturally deduced by further dividing the number of lanes. This concept can be described as continuous in space but discrete in time. However, practical traffic network system, particularly freeway network, does not have aerial camera to continuously monitor the traffic in real-time. Although dense point sensor system could approximate continuous measurement in space, the cost is prohibitive in general. In practice, point sensors such as inductive loop detectors are popularly used for traffic detection, which could continuously count vehicle number in traffic stream in real-time at the sensor locations. i.e. practical data provide information continuous in time but discontinuous in space. This is the gap between the state of traffic monitoring and the practical data requirement for real-time traffic control such as CRM and VSL.

Basically, the accuracy of measured flow from existing systems can be reasonably high, while the accuracy of distance mean speed estimation has a large variance. Hence, it is difficult to obtain relatively good estimation for density only based on the flow/speed/density relation with estimated flow and speed data from existing systems. The proposed method derives from the definition of density, and counting the vehicle number within specific segments correctly is the key for this method. With the development of vehicle telematics and other ITS technologies such as GPS (Global Positioning System) and VII (Vehicle-Infrastructure Integration), information from the vehicle itself could be a potentially valuable source to compensate for the deficiency of the point sensors. Reliable flow estimation from loop detector system and accurate vehicle location information from VII probe vehicles make it feasible to count the vehicle numbers correctly within specified segments once the market penetration of vehicles with VII is increasing to some level. This is the basic idea of the proposed methodology.

1.2 Congestion and Density

Why density estimation is important? It is critical to traffic control such as CRM and VSL. To avoid the local undesirable effect of ramp metering, it is necessary to estimate not only the average density, but also spatial-temporal density distribution. Particularly, it is necessary to know when traffic stream with a low or high density is passing the metered onramp. By knowing this in real-time, the ramp metering strategy can be designed so that it not only optimizes the traffic from a macroscopic level which is usually based on average density of each cell (section, or segment), but also injects vehicle into the freeway in right time so that the density distribution is approaching even. It is thus necessary to implement ramp metering strategy in two levels: (a) optimization in a macroscopic level based on average density of each cell along a stretch of freeway, which determines the optimal reference average density; (b) to select appropriate time slot based one density distribution in mainline to avoid imputing vehicles into traffic stream with high density.

1.3 Research Scope

For a stretch of freeway, if inductive loops are installed, vehicle count or volume between two adjacent loop stations can be determined once the onramp and off-ramp flow are known. Based on the vehicle count at initial time instance, vehicle count change and the distance between the two loop stations, an average density can be estimated. However, density variation at each distance point with respect to time cannot be estimated unless the loop stations are dense enough, as this is cost prohibitive for a freeway system. Inductive loop detector systems have been widely deployed, it makes better sense to fully adopt available infrastructure to achieve required traffic measurement. The VII vehicle with GPS provides a continuous measurement of speed and position of the vehicle within a moving window in real-time. This is a very good data source which complementary to the fixed loop data measurement.

This report considers real-time estimation of the density along a stretch of freeways based on (a) inductive loop stations which provide point measurement of traffic speed and vehicle count; (b) certain percentile of probe vehicles which provide moving window over distance for speed and position.

2 Literature Review

It is necessary to develop a ramp metering strategy to control density distribution along a stretch of freeway. If measurement of traffic is not accurate enough and the time for traffic injection through onramp is not proper, the imputed vehicle from onramp could cause congestion. This leads to a new conclusion: ramp metering needs the estimation of density distribution, not just average density or volume in cell (section, or segment). The injection should be at proper time instant – if there is a traffic stream with high density passing by, and ramp meter should not impute any vehicle into the mainline.

A method is discussed for estimating the number of vehicles on a section of a freeway from speed and flow measurements at the entrance and exit points of the section (Gazis 1971). The author proposed one real-time estimation method of traffic densities from time-series of flow and speed data. The method consists of obtaining true estimate of vehicle count at regular intervals, and then filtering random errors of these estimates by means of a sequential correction scheme. Periodically an estimation of vehicle travel times is needed from time-series speed data, and then a rough estimate of vehicle count is obtained from the inventory of car arrivals at the entrance of the section and the travel times of existing cars. The algorithm has been tested using data from three adjoining half-mile sections in the Lincoln Tunnel. This method relies on the accuracy of travel time estimation very much.

Another method bases its estimation on two easily obtained measurements from loop detector data: the flow rate and occupancy (Andrew, 1980). The occupancy of a detector during a particular time interval is the percent of that interval during which the detector signals the presence of vehicles. From measurements at neighboring detector stations, the number of vehicles, and therefore the density, on the segment between stations is tracked. The estimate is calculated using a Kalman filter. The detection of spatially inhomogeneous conditions is done using a Generalized Likelihood Ratio (GLR) event detection scheme. Both the filter and the GLR algorithm are simple discrete-time scalar equations. It is claimed that the unique point of the study is that it estimates the segment density accurately in all types of traffic conditions, and it does this by detecting spatially inhomogeneous traffic conditions and compensating the density estimate appropriately for the adverse effects of the inhomogeneities.

The traffic states were considered as Markov chain as time evolving and used linearized Cell Transmission Model based Kalman filter to provide the estimation of spatiotemporal density by Sun et al (2003). Three methods are evaluated based on traffic data on I-210. The criterion used is the relative error at fixed distance point where the measurement is available. For a stretch of freeways, the real-time estimation of the density at those measurement points was incorporated to enhance the estimation process.

Vehicle Re-identification technique was used to estimate in-lane flow and lane density by Coifman (2003). A method was developed for estimating density in a freeway lane between detector stations and measuring the net number of vehicles to enter or leave the lane. It relies on the accuracy of vehicle re-identification algorithm very much. The author tried to implement lane inflow estimation as well, and it was a good starting point to investigate the feasibility of estimating lane density only using loop station data. There was no dynamic traffic model involved in the algorithm.

Another way to determine traffic flow characteristics, including traffic density, was by definition directly (Ni, 2007). The following criteria need to be met considering to measure traffic flow characteristics: it involves only one type of sensor that is capable of determining density, flow and distance mean speed; the measurement should be accurate; it should preserve the fundamental relationship among density, flow and distance mean speed; and it should be compatible with ITS application. The paper tackled the challenge by formulating a method called the n-t method, the inputs of this method were time-stamped traffic count. Further, it was assumed that a point sensor was able to accurately timestamp the passage of each vehicle at the sensor location. It was feasible to construct cumulative curves of vehicle counts in the cumulative count-time (n-t) domain. The method required that no vehicle was lost or gained within any given region, traffic should be counted and time-stamped accurately, the roadway section under analysis was empty at the initial time interval; and a mechanism was needed to eliminate error accumulation due to the use of cumulative vehicle counts.

Estimation can also be separated from control if the system state can be directly estimated from sensor data through filtering and/or fusion. In this approach, the sensor measurement and state estimation are independent form the control process. The approach usually does not predict the system state and thus traffic model is not involved in the estimation.

3 Method for Density Estimation

Although, video camera (as range sensor) data are good for density estimation, they are usually unavailable where they are needed. Most of existing traffic detection sensors are inductive loops (as point sensors) which can only generate vehicle count or volume at sensor locations.

Compared with single loop detector, dual loop detector can provide more accurate information which could help in estimating density. Assuming x_i and x_{i+1} are two consecutive loop detector stations which can determine one segment $\left(x_i, x_{i+1}\right)$, dual loop detector can provide spot based speed and vehicle count (equivalent to flow or volume) over the station:

i : detector station index

$v\left(x,k\right)$: spot-based speed at location x during time interval k , mph

$q\left(x,k\right)$: spot-based flow at location x during time interval k , vphpl

$v\left(x_i, x_{i+1}, k\right)$: segment-based speed at the section between location x_i and x_{i+1} during time interval k , mph

$q\left(x_i, x_{i+1}, k\right)$: average flow at the section between location x_i and x_{i+1} during time interval k , vphpl

$\rho\left(x_i, x_{i+1}, k\right)$: average density at the section between location x_i and x_{i+1} during time interval k , veh/mile/lane

Generally, if the distance between two consecutive loop stations is short, say less than 0.5~1 mile, we may approximate the distance mean speed as the follows:

$$v\left(x_i, x_{i+1}, k\right) = \frac{1}{2}\left(v\left(x_i, k\right) + v\left(x_{i+1}, k\right)\right)$$

As for average flow along the section $\left(x_i, x_{i+1}\right)$, it can be approximated by:

$$q\left(x_i, x_{i+1}, k\right) = \frac{1}{2}\left(q\left(x_i, k\right) + q\left(x_{i+1}, k\right)\right)$$

According to the relation among flow, speed and density, the average density between the section $\left(x_i, x_{i+1}\right)$ can be roughly calculated by:

$$\rho\left(x_i, x_{i+1}, k\right) = \frac{q\left(x_i, x_{i+1}, k\right)}{v\left(x_i, x_{i+1}, k\right)}$$

In this paper the above procedure is termed as Method 1 dedicated to using only loop data for density estimation. However, this is not adequate for the estimation of density distribution between two terminal loop stations along a stretch of freeway. It cannot guarantee good results in case there exists on-ramp or off-ramp at the section, or the length of the section $\left(x_i, x_{i+1}\right)$ is rather long. Further, the definition of section $\left(x_i, x_{i+1}\right)$ is determined by the loop detector locations, and the section definition may be not compatible or appropriate for traffic control.

3.1 Spatiotemporal Density Estimation

With the help of VII probe vehicles, the density distribution estimation between two loop stations or in the section between one loop station and one specific location (the location of VII probe vehicle) at downstream would be feasible. Meanwhile, the definition of section can be more flexible. For example, we can even dynamically define sections based on the need for real-time traffic control.

Existing traffic monitoring systems are based on the vehicle count and occupancy from loop station, which is distributed every 0.5 to 2 miles along freeway. Dual loop station can provide reasonably accurate vehicle speed and vehicle count at one point station. However, this is not adequate for the estimation of density distribution between the two terminal stations along a section of freeway.

With the help of VII vehicles, the density distribution estimation between two loop stations would be feasible. This can be sketched as follows:

- Dual loop station provides distance mean speed or time mean speed over the station;

  speed distribution $v\left(x_1, t\right)$ and $v\left(x_2, t\right)$

  density distribution $\rho\left(x_1, t\right)$ and $\rho\left(x_2, t\right)$

- Based on the speed and density estimation at the two loop stations and the traffic states (conditions), and the shockwave propagation speed, it is possible to provide an estimation of

  speed distribution $v(x,t)$

  density distribution $\rho(x,t)$

If VII vehicles have certain level of market penetration, the information from those vehicles can be synchronized with loop data for real-time continuously spatial and temporal estimation of speed and density along a stretch of freeways. This can be conducted as follows:

(1). Estimation of speed and density at point stations;

(2). Synchronizing/fusing the VII/probe vehicle speed and the loop detector detected speed;

(3). Compute density between upstream loop detector station and the current location of VII probe vehicle based on vehicle count and distance between VII probe vehicle and upstream loop station;

(4). Conduct step (3) for each VII probe vehicle, and then interpolate the density around all the VII probe vehicles to obtain the estimation of density continuous in both time and distance.

Since the spatiotemporal characteristics of density is the output of dynamical system with involves both speed and density, and possibly with on-ramp flow. The estimation of the density must involve some dynamic evolving, which means that a model based approach should be taken. With the model as the basis, updated information at the discrete loop stations and the information from probe vehicles should be taken into consideration.

3.2 Assumptions of the Proposed Method

Figure 1 shows the principle of the proposed method with assumptions:

(1). There is a loop detector at x_i, but not necessarily at x_{i+1};

(2). If there exists on-ramp and/or off-ramp between location x_i and x_{i+1}, detector is assumed to be installed for counting on-ramp and/or off-ramp vehicle;

(3). The timestamps of loop detector data and GPS from VII vehicle are synchronized;

(4). If there exist vehicles bypassing VII probe vehicle p during time interval k, the probe vehicle p bypass the same number of vehicles during the same time interval. This is naturally true if lane changing rate is low and can be skipped.

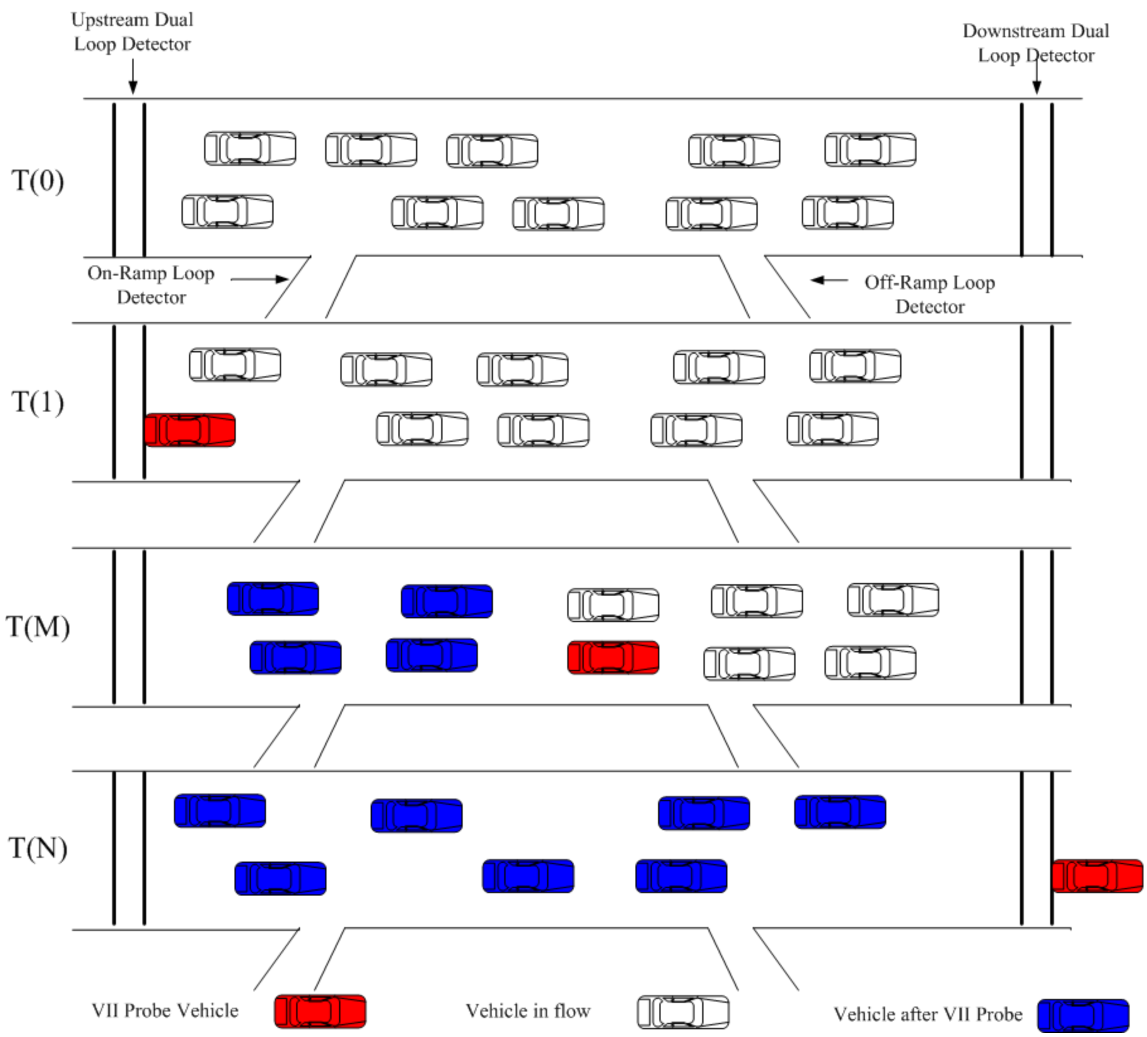

Figure 1. Density Estimation with VII Probe Vehicle and Loop Detector Data

3.3 Notations of the Proposed Method

k : index of time interval, $k = 0,1,2,......K$

T : sample time interval for average density estimation, T can be 10 sec, 20 sec, 60 sec, 300 sec, etc., second

T_u : time interval for VII probe vehicle location update, T_u can be 1 sec, 10 sec, 20 sec, etc., second

j : time index for sampling

$t = jT_u$: timestamp of vehicle arrival at specified location

$L(x, y)$: distance between location x and y , mile

$L(x_i, x_{i+1})$: distance between location x_i and x_{i+1} which define a cell, mile

$L(x_i, p)$: distance between location x_i and VII probe vehicle p , mile

$LN(x_i, x_{i+1})$: number of lanes between location x_i and x_{i+1}

$N(x,t)$: accumulative vehicle count from loop station at location x from time 0 to t

$N_r(x_i, x_{i+1}, t)$: accumulative vehicle count from on-ramp loop station between x_i and x_{i+1} from time 0 to t

$N_s(x_i, x_{i+1}, t)$: accumulative vehicle count from off-ramp loop station between x_i and x_{i+1} from time 0 to t

$t(x, p)$: arrival time for VII probe vehicle p at location x

$\rho(x_i, x_{i+1}, t, p)$: calculated density from vehicle p at the section between location x_i and x_{i+1} during time interval k , veh/mile

$\rho(x_i, x_{i+1}, k)$: average density at the section between location x_i and x_{i+1} during time interval k , veh/mile

$\hat{\rho}(x_i, x_{i+1}, k)$: estimated density at the section between location x_i and x_{i+1} during time interval k , veh/mile

3.4 Procedures of the Proposed Method

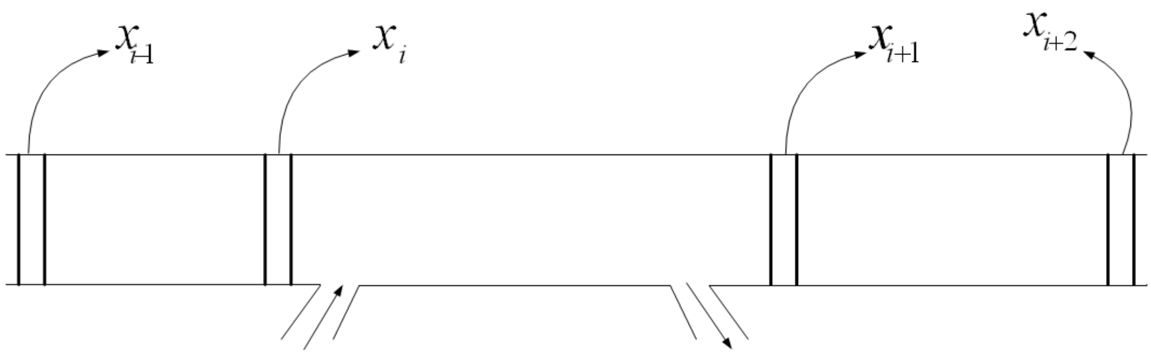

Figure 2. Three consecutive cells

Considering three typical consecutive cells: $\left[x_{i-1}, x_i\right), \left[x_i, x_{i+1}\right)$ and $\left[x_{i+1}, x_{i+2}\right)$, the general procedure of estimating average density can be formulated as following:

Step 1: Initialization: at the first time interval $k = 0$, if there exist one valid sample for density estimation, go to step 2 directly; if no, set $\rho\left(x_i, x_{i+1}, 0\right)$ as one default density;

Step 2: During the time interval $k = 1, 2, ...K$, $t \in [(k-1)\times T, kT)$, sampling for every T_u second if one vehicle passed the location x_1 already and is at the location y (y may be between location x_i and x_{i+1} or may be at the downstream site of location x_{i+1}) , the following two inequalities hold $\dfrac{L(x_i, y)}{L(x_i, x_{i+1})} \geq 80\%$ and $\dfrac{L(y, x_{i+1})}{L(x_i, x_{i+1})} \leq 20\%$, then calculate the measured density at the time instant t from vehicle p at the section between location x_i and x_{i+1} as follows.

Main lane vehicle count change:

$$\Delta N = N(x_i, t(x_{i+1}, p)) - N(x_i, t(x_i, p))$$

On-ramp vehicle count change:

$$\Delta N_r = N_r(x_i, x_{i+1}, t(x_{i+1}, p)) - N_r(x_i, x_{i+1}, t(x_i, p))$$

Off-ramp vehicle count change:

$$\Delta N_s = N_s(x_i, x_{i+1}, t(x_{i+1}, p)) - N_s(x_i, x_{i+1}, t(x_i, p))$$

$$\rho(x_i, x_{i+1}, t, p) = \frac{\Delta N + \Delta N_r - \Delta N_s}{L(x_i, x_{i+1}) \times LN(x_i, x_{i+1})}$$

Otherwise, $\rho(x_i, x_{i+1}, t, p) = 0$

Step 3: Detecting: if there are M_p calculated density from same probe vehicles p during the time interval k, then $\rho(x_i, x_{i+1}, k, p) = \frac{1}{M_p}\sum \rho(x_i, x_{i+1}, t, p)$ for all

$$\rho(x_i, x_{i+1}, t, p) \neq 0$$

Step 4: At the end of the time interval k, averaging:

a. For one cell determined by location x_1 and x_2 during time interval k, if there are P valid VII probe vehicles for sampling, then

$$\rho(x_i, x_{i+1}, k) = \frac{1}{P}\sum_{p=1}^{P} \rho(x_i, x_{i+1}, k, p)$$

b. If there is no valid VII probe vehicles for sampling for the cell, then

$$\rho(x_i, x_{i+1}, k) = \alpha_1 \times \rho(x_{i-1}, x_i, k-1) + \alpha_2 \times \rho(x_{i-1}, x_i, k) + \alpha_3 \times \rho(x_i, x_{i+1}, k-1)$$
$$+\alpha_4 \times \rho(x_{i+1}, x_{i+2}, k-1) + \alpha_5 \times \rho(x_{i+1}, x_{i+2}, k)$$

3.5 Parameter Identification of the Proposed Method

Ground truth data (exact vehicle count for each time interval) and estimated vehicle count are needed for further parameter identification. Currently, we can apply the following proposed strategy because of insufficient data source.

Case 1: If there exist valid VII probe vehicles from upstream (x_{i-1}, x_i) and there is no valid probe vehicles from downstream (x_{i+1}, x_{i+2})

$\alpha_1 = 0.3, \alpha_2 = 0.3, \alpha_3 = 0.4, \alpha_4 = 0, \alpha_5 = 0$

Case 2: If there is no valid VII probe vehicles from upstream (x_{i-1}, x_i) and there exist valid VII probe vehicles from downstream (x_{i+1}, x_{i+2})

$\alpha_1 = 0.3, \alpha_2 = 0, \alpha_3 = 0.4, \alpha_4 = 0, \alpha_5 = 0.3$

Case 3: If there is no valid VII probe vehicle from either upstream (x_{i-1}, x_i) or downstream (x_{i+1}, x_{i+2})

$\alpha_1 = 0.3, \alpha_2 = 0, \alpha_3 = 0.4, \alpha_4 = 0.3, \alpha_5 = 0$

Case 4: If there exist valid VII probe vehicles from both upstream (x_{i-1}, x_i) and downstream (x_{i+1}, x_{i+2})

$\alpha_1 = 0, \alpha_2 = 0.3, \alpha_3 = 0.4, \alpha_4 = 0, \alpha_5 = 0.3$

In this paper the above procedure is termed as Method 2 using both loop detector and probe vehicle data as input for density estimation.

4 Method Validation

4.1 BHL Data System

To validate the proposed method, the algorithm in Section 3 has been implemented and the practical data from the Berkeley Highway Laboratory (BHL) system have been used. BHL is a test site which covers 2.7 miles of I-80 immediately east of the San Francisco-Oakland Bay Bridge in California, which can be seen from Figure 3. BHL provides event data on individual vehicle actuations, accurate to 1/60th of a second (60 Hz data). Most other loop detector systems collect only aggregated data over periods of 20 seconds or even longer (1 minute, or 5 minutes).

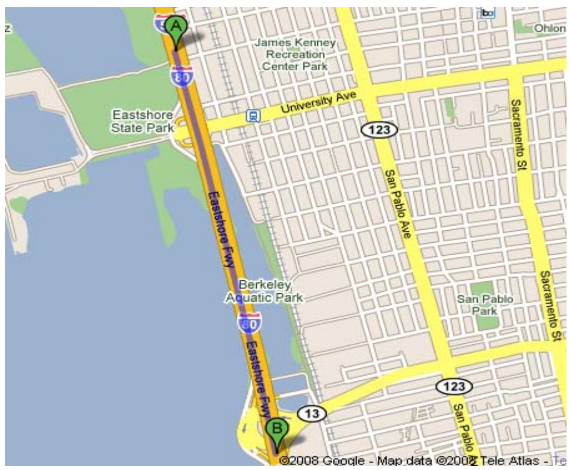

Figure 3. Studied Freeway Stretch – I-80

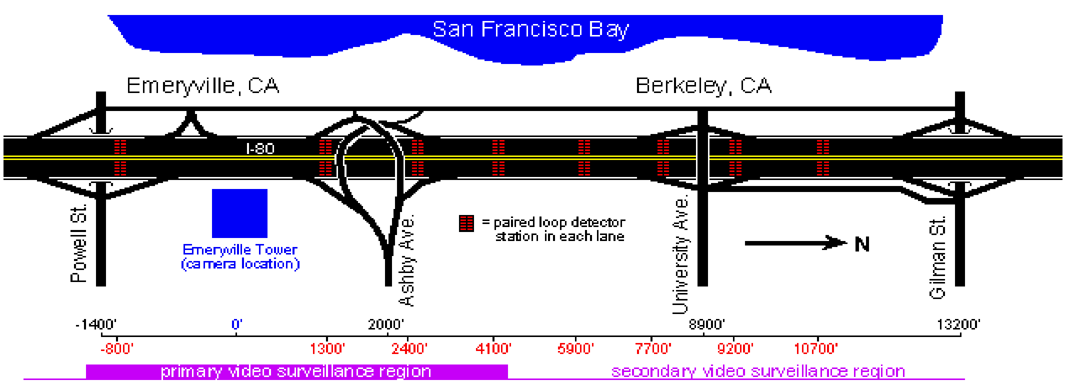

Figure 4. Monitored Freeway Stretch by BHL Data

The format of the raw sorted loop detector data can be seen as the following:

Table 1 Specification of the raw sorted BHL loop data

TimeStamp_1	TimeStamp_2	170 Controller
1218697274000	1218697273000	27150973F
1218697275000	1218697274000	271509830
1218697276000	1218697275000	271509972C9FEAADACCCFDCDEE0DEAA3F
1218697277000	1218697276000	271509A0DD26928A5896708F1910A6BB91D68D8B76
1218697278000	1218697277000	271509B188EF048
1218697279000	1218697278000	271509C2F13105D17C
1218697280000	1218697279000	271509D4C
1218697281000	1218697280000	271509E4D
1218697282000	1218697281000	271509F4E
1218697283000	1218697282000	27150A02705315365D08168378D1B1739
1218697284000	1218697283000	27150A1B34B3EDD4E35E3DE5E40
1218697285000	1218697284000	27150A21761D521D47749679743
1218697286000	1218697290000	27150A39F4CF5D1479
1218697287000	1218697290000	27150A4115D6B41
1218697288000	1218697290000	27150A516814B44805
1218697289000	1218697290000	27150A646
1218697290000	1218697291000	27150A747
1218697291000	1218697291000	27150A848
1218697292000	1218697291000	27150A926A56954A84943
1218697293000	1218697292000	27150AAE2771
1218697294000	1218697293000	27150AB1C72204C07EBAE8B0BDC803

The first field is epoch time in milliseconds, and generally is expressed as elapsed time since midnight (beginning of the calendar day) GMT on January 1, 1970. The second field is currently identical to the first field. The third filed represents the raw data from the 170 controller (refer to: http://www.bhl.calccit.org/). The exact location including longitude and latitude of each loop detector is known to us, and it makes it feasible to know the distance between each loop detector and any vehicle of interest and known location.

In the real world, there are no real-time VII probe vehicle data in the studied area. To collect the GPS data to simulate as VII prove vehicle data, we drove two Infinite SUV vehicles five rounds along the studied freeway segment and the data were recorded at 20Hz. In this study, both BHL loop data and the GPS data from test driving are used. We compared two types of density estimation output: estimated density from loop detector data, and estimated density from the proposed method which uses both loop detector data and the GPS. The BHL loop detector data used for this study includes both East-bound

(EB) and West-bound (WB) direction of I-80 on August 14 2008. The format of the GPS data from test driving can be seen as the following:

Table 2 Specification of GPS data from test driving

Longitude	Latitude	Timestamp
122.33342650000002	37.915342833333334	08:56:41.997
122.3334266666667	37.9153426666667	08:56:42.047
122.3334266666667	37.9153426666667	08:56:42.097
122.3334266666667	37.9153426666667	08:56:42.147
122.3334266666667	37.9153426666667	08:56:42.197
122.333427	37.9153426666667	08:56:42.247
122.333427	37.9153426666667	08:56:42.297
122.333427	37.9153426666667	08:56:42.347
122.333427	37.9153426666667	08:56:42.397
122.33342733333332	37.915342499999994	08:56:42.447
122.33342733333332	37.915342499999994	08:56:42.497
122.33342733333332	37.915342499999994	08:56:42.547
122.33342733333332	37.915342499999994	08:56:42.597

The first two fields show the longitude and latitude of the corresponding vehicle at each time instance. The third field represents the timestamp in which the record is generated, and the accuracy of timestamp is millisecond.

4.2 Method Validation with BHL and Field Collected Data

The scenarios selected span a wide variety of traffic conditions (free flow, moderate flow), and density has been estimated every 20 seconds. The density estimation results based on two methods are shown in Figure 6 (I-80 WB). The green line is the estimation from two terminal dual loop stations, and the blue spot is from VII probe vehicle and dual loop stations. The length of the section is 334 meters (0.21 mile) with 5 lanes. It is observed from Figure 4 that at time interval around 9:08AM, 9:40AM, and 10:08AM, the estimated results from two methods are quite similar. At 9:22AM, it is still in free flow speed (~ 60 mph), while the flow demand from upstream section is increasing, the density estimation from Method 2 is more sensitive to the change, compared to the Method 1. At 9:53AM, the speed from loop detector has been dropped to around 20 mph for only one time interval for some reason, and it is likely caused by some disturbance rather than actual traffic stream, and then speed went back to moderate flow about 40~50 mph. The estimated density from Method 1 is around 70 veh/mile, while the one from Method 2 is 28 veh/mile. The distance mean speed estimation in Method 1 has over restricted assumptions which do not hold all the time, and speed estimation cannot guarantee to be reasonable over time. Hence, the density estimation from Method 1 has large errors at times. Because the vehicle count data is relatively good, the estimate from VII probe vehicle and loop detector is able to track the transient behavior of density accurately.

Figure 5. Station 6 – Station 7 @ I-80 West, 08/14/2008

Table 3 Calculated density based on Loop Detector and VII Probe Data

Lane	Station_6	Station_7	St_6_Speed	St_6_Count_1	St_6_Count_2	Density
5	432	448	38	135	139	19
4	1392	1410	63	459	461	10
3	2349	2364	34	852	856	19
2	3258	3272	44	1533	1540	34
1	4078	4092	58	1919	1924	24
5	426	443	39	132	138	29
4	1387	1404	63	456	461	24
3	2345	2360	36	851	854	14
2	3244	3258	49	1525	1533	38
1	4070	4084	54.3	1913	1920	34

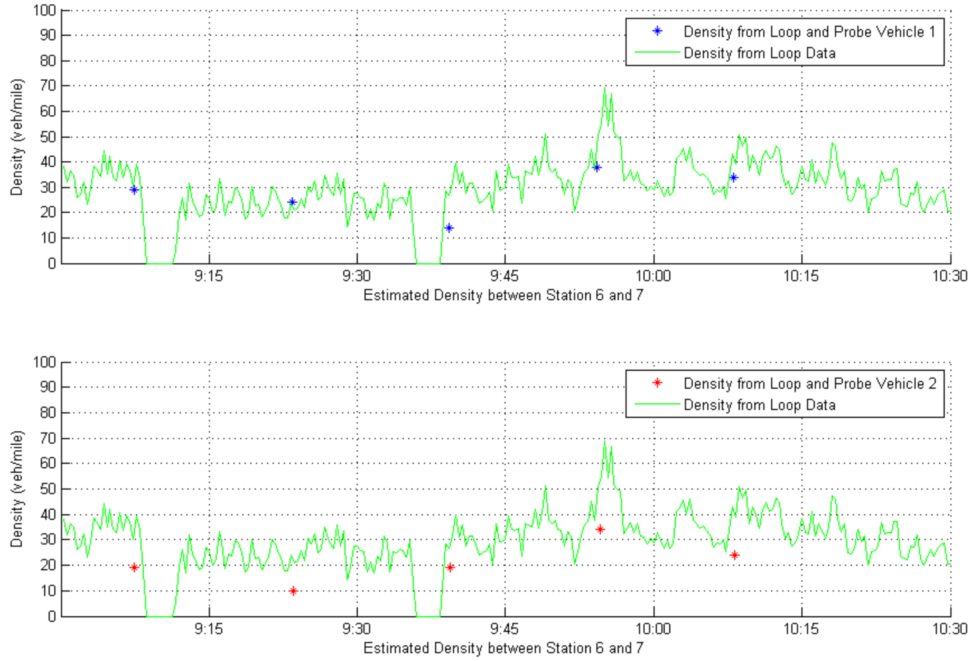

Figure 6. Segment Density Results Comparison between two methods

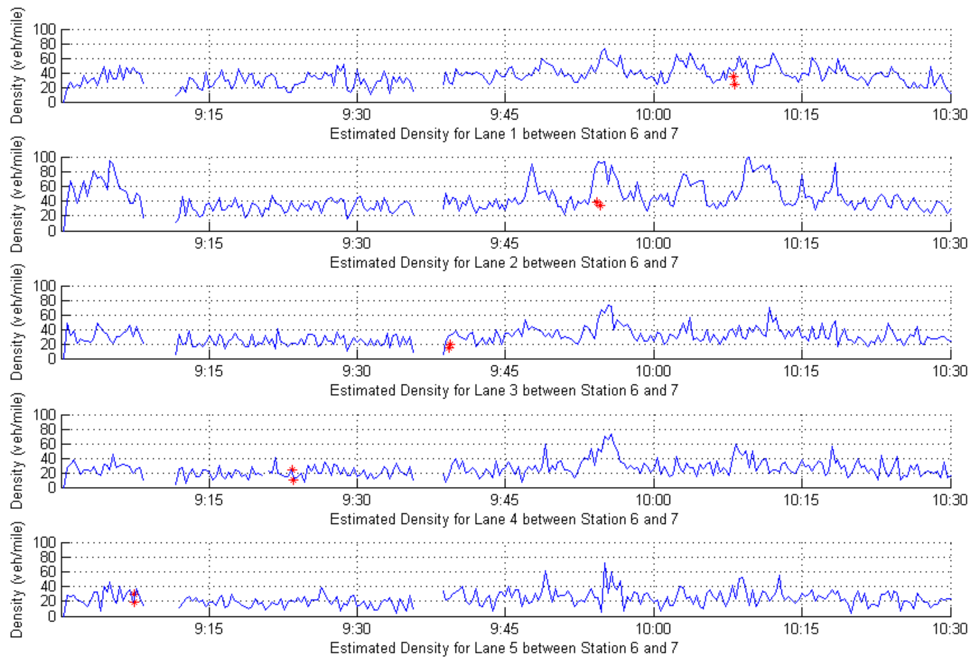

Figure 7. Lane Density Results Comparison between two methods

Figure 7 represents the comparison results of density estimation for each lane individually. The blue line is the estimation from two terminal dual loop stations, and the red spot is from VII probe vehicle and dual loop stations. The five figures represent the five individual lanes from the top to bottom respectively. Similarly, at 9:08AM and 9:20AM the two methods can generate similar results. At 9:53AM, for the same reason, Method 2 is not affected by the accuracy of speed estimation, and is able to track the vehicle count change more accurately. While at 9:40AM and 10:08AM, it is difficult to judge which one has better estimation results, as we don't have ground truth data to compare. In this case, the error of Method 2 comes from the uncertainty of lane change flow. As the length of section is a bit short, the small error of vehicle count can lead to big error of density estimation. This is the limitation of the proposed method applied into lane specific density estimation if the market penetration of VII vehicle is low. If the market penetration is high, there is no such as problem since individual vehicle behavior including lane changing is known. This is the new trend since lane specific density estimation is needed for traffic control such as CRM and VSL control.

4.3 Method Validation with NGSIM data

In 2005, data on individual vehicle trajectories were collected and made available under the Next Generation Simulation (NGSIM) project. The NGSIM data provide a unique opportunity to verify and validate the developing microscopic and macroscopic traffic models. The NGSIM freeway data consists of vehicle trajectories on two test sites: I-80 and US-101 test sections. The I-80 (BHL covered area) test section is chosen for this study, and it is a 0.3125 mile (0.5 km) 6-lane freeway weaving section. The processed data includes 45 minutes of vehicle trajectories in transition period (4:00-4:15PM) and congestion period (5:00-5:30 PM).

In the NGSIM trajectory data, each individual vehicle's position and timestamp were captured every 1/10 second, which can be used as the ground truth of density estimation. In using the NGSIM data, we take each vehicle as virtual VII probe vehicle and take the starting and ending points of the section as virtual loop detector stations. In order to validate the proposed algorithm, with Method 1 (using only loop detector data) and Method 2 (using both virtual loop detector and virtual VII probe vehicle data), density for the section has been estimated every 20 seconds and every 60 seconds for the same covered time period with NGSIM trajectory data.

The density estimation results based on two methods and ground truth from NGSIM video data are shown in Figure 9 to Figure 14 (I-80 EB). The blue lines in the figures show the density estimation from two terminal dual loop stations, the green lines show the estimated density from virtual loop detector data and virtual VII probe vehicle data. The length of the section is 1650 feet (0.3125 mile, 502.9 meters) with 6 lanes, and three consecutive fifteen minutes data have been analyzed for 04/13/2005. It is transition time period (from free flow to congestion) at 4:00PM-4:15PM, and high congestion period at 5:00PM – 5:30PM.

It is observed from Figure 9 and Figure 10 that the two methods can generate similar estimations from 4:00PM to 4:07PM, while the Method 2 can catch the density dynamic change better than the Method 1 considering the results from 4:07PM to 4:15PM. Similar conclusions can be draw from Figure 11 to Figure 14, the Method 1 has large error than the Method 2 has, whether for 20 second dataset, or for 60 second one. Table 4 detailed the Mean Square Error (MSE) and Root Mean Square Error (RMSE) for different scenarios, and confirmed the conclusions shown from the figures.

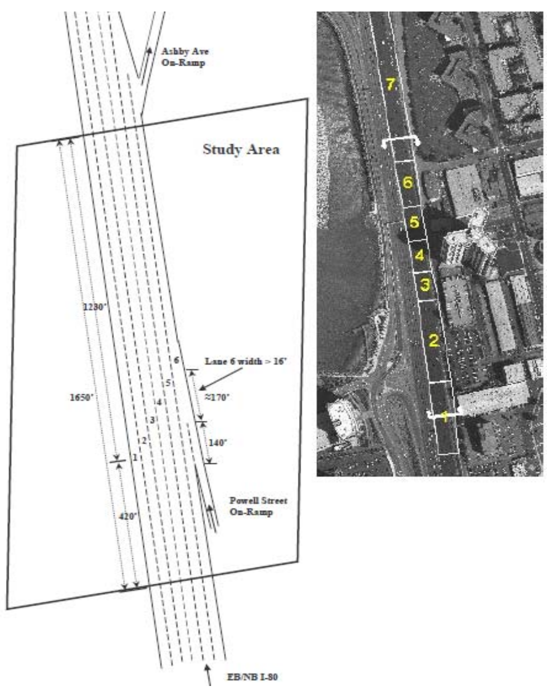

Figure 8. NGSIM Study Area and Camera Coverage

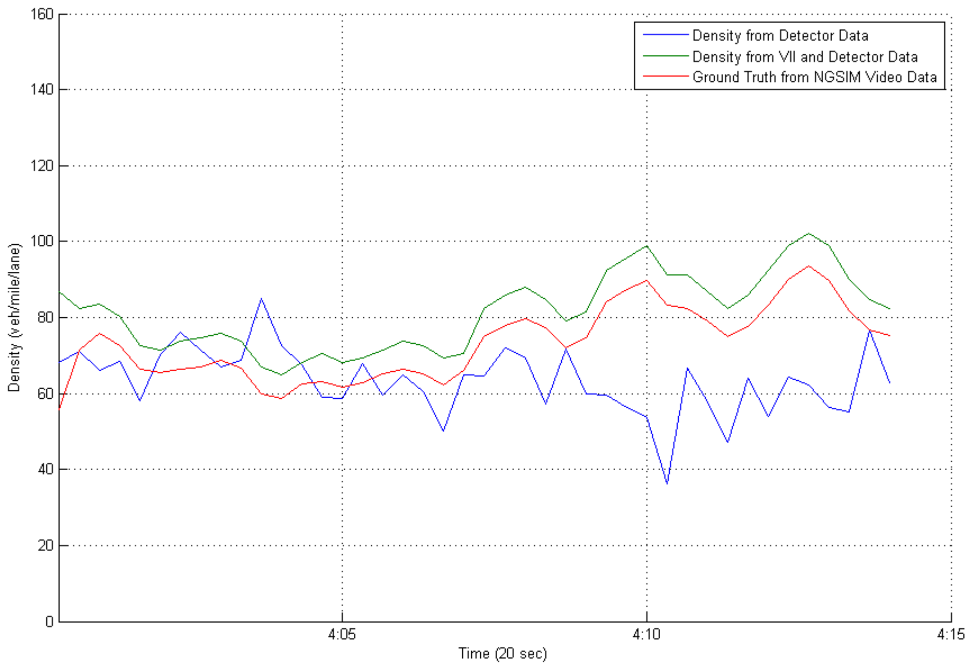

Figure 9. Results Comparison for 20-sec Dataset at 4:00-4:15 PM on 04/13/2005

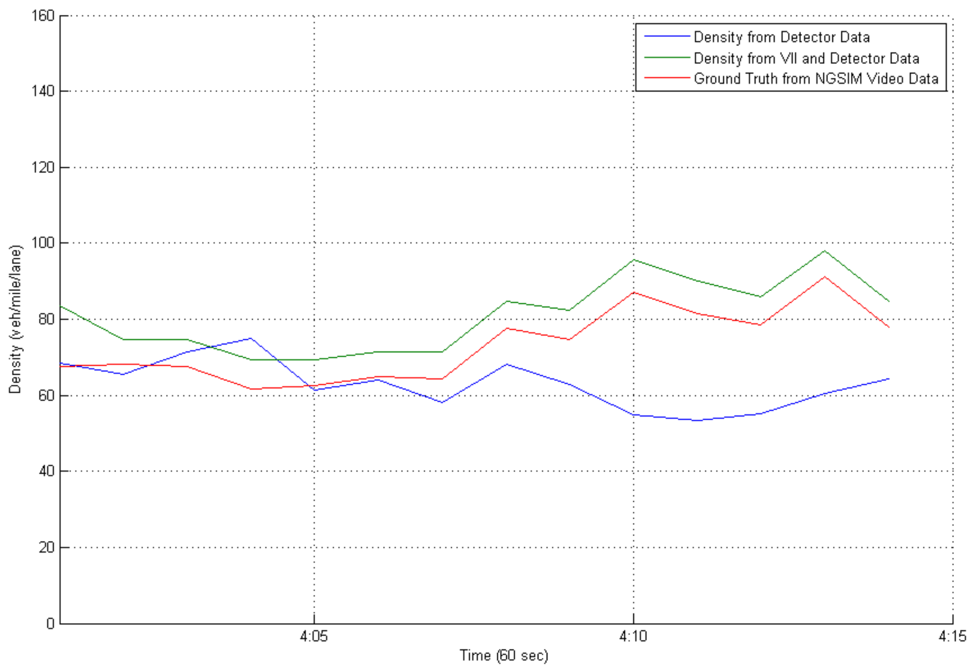

Figure 10. Results Comparison for 60-sec Dataset at 4:00-4:15 PM on 04/13/2005

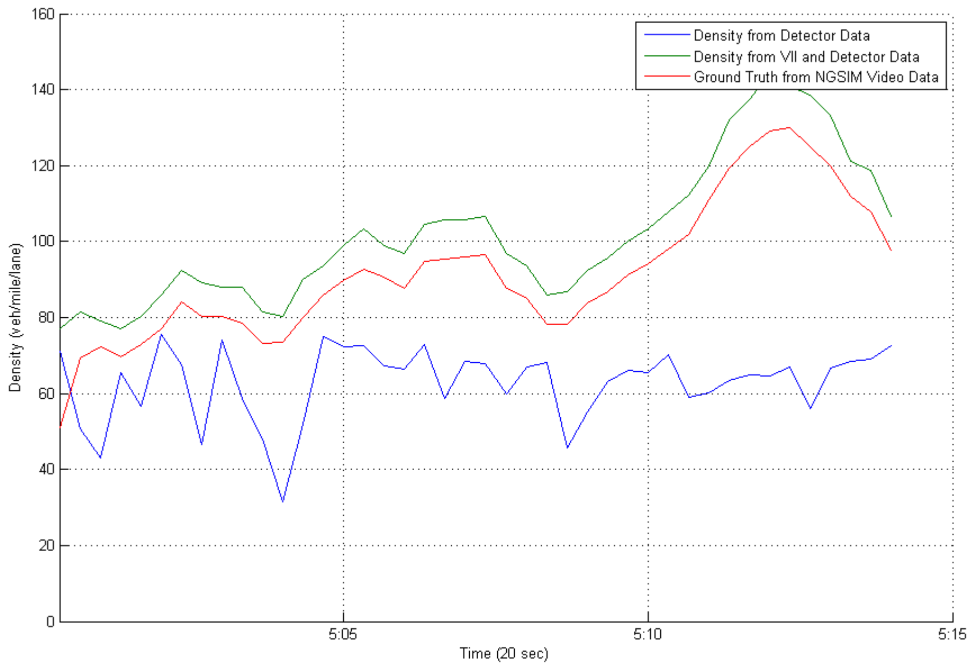

Figure 11. Results Comparison for 20-sec Dataset at 5:00-5:15 PM on 04/13/2005

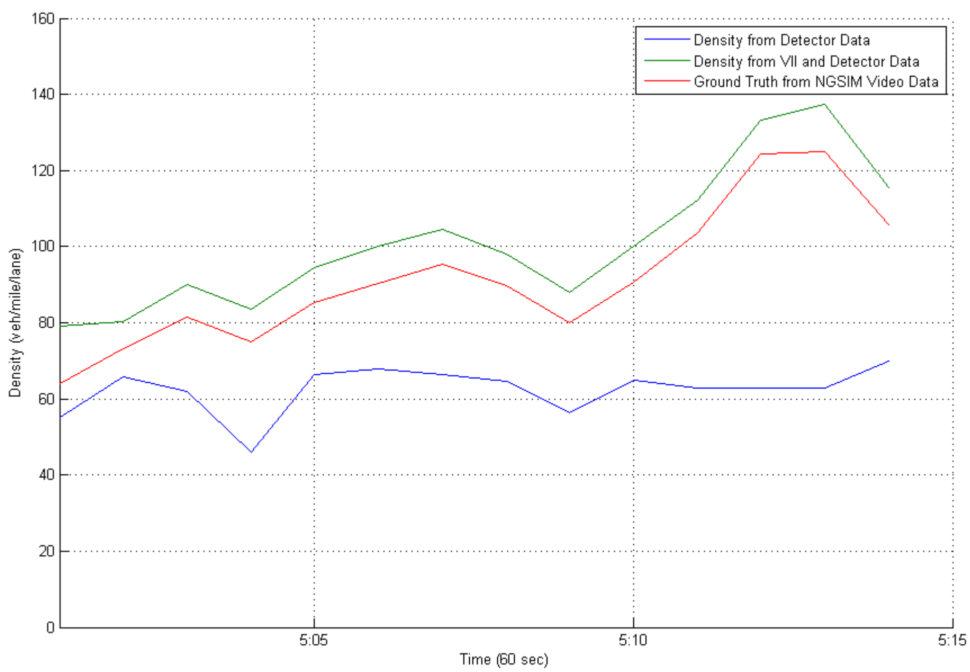

Figure 12. Results Comparison for 60-sec Dataset at 5:00-5:15 PM on 04/13/2005

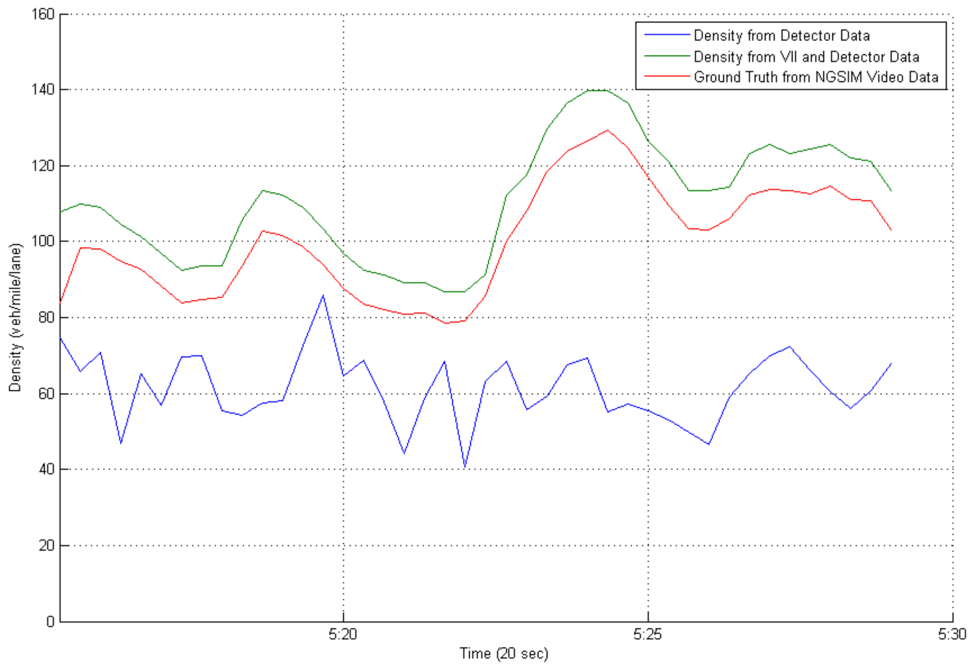

Figure 13. Results Comparison for 20-sec Dataset at 5:15-5:30 PM on 04/13/2005

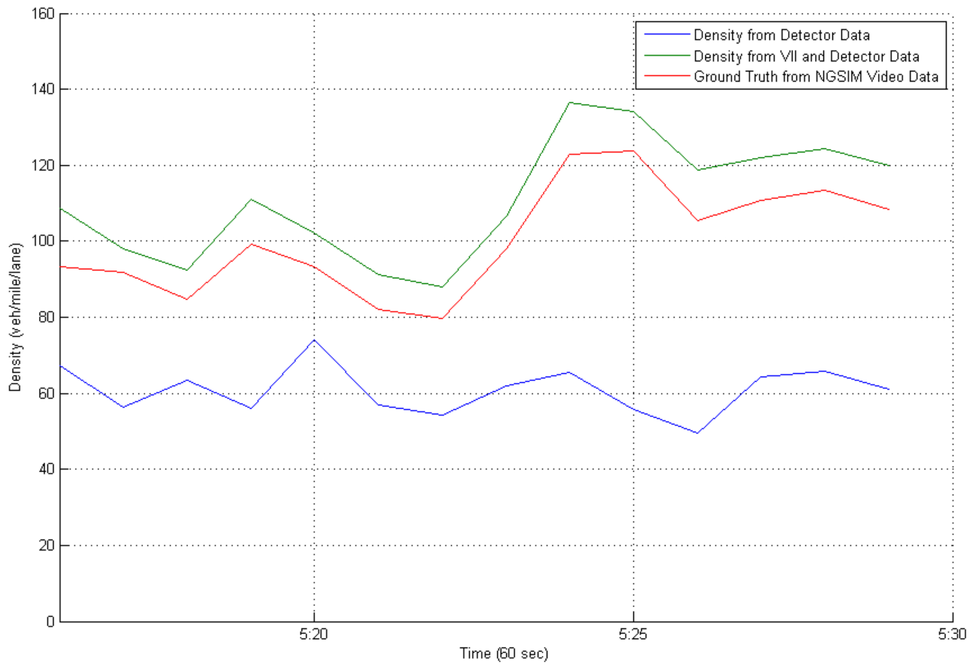

Figure 14. Results Comparison for 60-sec Dataset at 5:15-5:30 PM on 04/13/2005

Table 4 MSE and RMSE Comparison Between Two Methods

		Time Interval = 20 seconds		Time Interval = 60 seconds	
		Detector Method	VII+Detector Method	Detector Method	VII+Detector Method
4:00 - 4:15	MSE	325.10	13.96	284.00	5.80
	RMSE	18.03	3.74	16.85	2.41
5:00 - 5:15	MSE	1165.90	10.35	1107.40	5.85
	RMSE	34.15	3.25	33.28	2.42
5:15 - 5:30	MSE	1791.80	5.98	1776.80	4.17
	RMSE	42.33	2.45	42.15	2.04

5 Conclusion and Remarks

In this study we tried a new method to integrate the advantages of loop detector and VII probe vehicle information to estimate density for pre-defined freeway segments at a specified time interval in real-time.

More reliable flow estimation from loop detector system and more accurate vehicle location information from VII probe vehicles make it feasible to count the number of vehicles correctly in the time-varying segment bounded by upstream loop station and the probe vehicle. The proposed method derives from the definition of density, and counting the vehicle number within the time-varying freeway segments correctly is the key for this method. With the proposed method applied, segments for traffic network can be defined more freely, and more accurate density estimation can be obtained as well. Experimental results show that the proposed method can generate more reasonable density estimated results – close to the density estimated from NGSIM data.

There is no on ramp and off ramp data currently available in BHL loop detector system. However, the method is still valid once the ramp data are available. In fact, the proposed method should have much higher accuracy if ramp data are available.

Lane specific density estimation may be needed in the practical real-time traffic control system. If the proposed method combining loop detector and VII probe vehicle is applied, the high accuracy of GPS units are needed on vehicle so that VII vehicles can be mapped into each individual lane.

References

BHL Homepage: http://www.bhl.calccit.org/

NGSIM Homepage: http://ngsim.fhwa.dot.gov/

Greenshields, B.D.(1933) The Photographic Method of studying Traffic Behaviour; Proceedings of the 13th Annual Meeting of the Highway Research Board.

Greenshields, B.D.(1935) A study of highway capacity. Proceedings Highway Research Record, Washington Volume 14, pp. 448-477.

Coifman, B. (2003) Estimating Density and Lane Inflow on a Freeway Segment, *Transportation Research: Part A*, vol 37, no 8, 2003, pp 689-701.

Muñoz, L., X. Sun, R. Horowitz, and L. Alvarez. (2003) Traffic Density Estimation with the Cell Transmission Model. *Proceedings of the 2003 American Control Conference*, Denver, Colorado, USA, June 2003, pp. 3750-3755.

Muñoz, L. (2004a) Macroscopic Modeling and Identification of Freeway Traffic Flow. Ph.D. Dissertation, University of California, Berkeley, 2004.

Muñoz, L., X. Sun, D. Sun, G. Gomes, and R. Horowitz. (2004b) Methodological Calibration of the Cell Transmission Model. *Proceedings of the 2004 American Control Conference*, Boston, Massachusetts, USA, June 30 – July 2, 2004, pp. 798-803.

Sun, X., Muñoz, L., Horowitz, R., and Alvarez, L. (2003) Highway traffic state estimation using Improved Mixture Kalman Filters for effective ramp metering control . *Proceedings of the IEEE CDC*, Maui Hawaii, USA, Dec. 9-12, 2003, pp.6333-6338.

Sun, X., Muñoz, L., and Horowitz, R. (2004) Mixture Kalman Filter Based Highway Congestion Mode and Vehicle Density Estimator and its Application. In *Proceedings of the 2004 American Control Conference*, Boston, Massachusetts, USA, June 30 – July 2, 2004, pp. 2098-2103.

Gazis D. C., and Knapp C. H. (1971) On-Line Estimation of Traffic Densities from time-Series of Flow and Speed Data, *Transportation Science,* Volume 5, Number 3, August, 1971

Papageorgiou M., and Vigos G. (2008) Relating time-occupancy measurements to space-occupancy and link vehicle count, *Transportation Research Part C*, 16 (2008) 1-17

Ni D., (2007) Determining Traffic-Flow Characteristics by Definition for Application in

ITS, *IEEE Transactions on Intelligent Transportation Systems*, Vol.8, No.2, June 2007

Nahi N. E. and Trivedi A. N. (1973) Recursive Estimation of Traffic Variables: Section Density and Average Speed, *Transportation Science*, Vol. 7, No. 3, August 1973, pp. 269-286

Andrew Kurkjian, Stanley B. Gershwin, Paul K. Houpt, Alan S. Willsky, E. Y. Chow, C. S. Greene (1980) Estimation of Roadway Traffic Density on Freeways Using Presence Detector Data, Transportation Science, Vol. 14, No. 3, August 1980, pp. 232-261

Papageorgiou, M. (1987), Macroscopic Modeling of Traffic Flow on the Boulevard Pe´riphe´rique de Paris. Rapport de Contrat INRETS (Institute National de Recherche sur les Transports et leur Securite´), Arcueil, France.

Bhouri, N., Hadj-Salem, H., Papageorgiou, M., Blosseville, J.M.(1988). Estimation of traffic density on motorways. Traffic Engineering and Control 29, 579–583.

Cassidy, M.J., Coifman, B. (1997). Relation among average speed, flow, and density and analogous relation between density and occupancy. Transportation Research Record 1591, 1–6.

Edie, L.C. (1965). Discussion of traffic stream measurements and definitions. In: Proceedings of Second International Symposium on the Theory of Traffic Flow, Paris, France, pp. 139–154.

Appendix

Matlab Code for Density Estimation

```
%% Calculate the mean Occupancy, Speed of segment of interest
%% function Est_State(date_str,up_st_str,dn_st_str)

date_str = '2008-08-14';
st_str = '7';
up_st_str = '6';    % upstream loop station #
dn_st_str = '7';    % downstream loop station #
st_dir = 'W';

Start_Hour = 9;

DeltaT_SampleTime = 20;  % unit is second
Total_SampleTime = 5400;
No_Interval = Total_SampleTime/DeltaT_SampleTime; % 5400 / 60 = 90, one
hour and half

N_state=4;
N_dat=100000;

%%%%%%%%%%%%%%% Begin %%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%% Lane 1 %%%%%%%%%%%%%%%%%

   fid_1=0;
   while fid_1<1                          %% Open file for reading
     [fid_1, message]=fopen(['st',st_str,'-',date_str,'-',st_dir,'-
1.dat'],'r');
     if (fopen(fid_1) == -1)
        disp(message);
     end
   end
   [A0,Count]=fscanf(fid_1,'%f %f %f %f',[N_state,N_dat]);
   A0=A0';
   fclose(fid_1);

%%%%%%%%%%%%%%% Lane 2 %%%%%%%%%%%%%%%%%

   fid_1=0;
   while fid_1<1                          %% Open file for reading
     [fid_1, message]=fopen(['st',st_str,'-',date_str,'-',st_dir,'-
2.dat'],'r');
     if (fopen(fid_1) == -1)
        disp(message);
     end
   end
   [A1,Count]=fscanf(fid_1,'%f %f %f %f',[N_state,N_dat]);
   A1=A1';
```

```
   fclose(fid_1);

%%%%%%%%%%%%%%% Lane 3 %%%%%%%%%%%%%%%%%

   fid_1=0;
   while fid_1<1                     %% Open file for reading
     [fid_1, message]=fopen(['st',st_str,'-',date_str,'-',st_dir,'-
3.dat'],'r');
     if (fopen(fid_1) == -1)
        disp(message);
     end
   end
   [A2,Count]=fscanf(fid_1,'%f %f %f %f',[N_state,N_dat]);
   A2=A2';
   fclose(fid_1);

%%%%%%%%%%%%%%% Lane 4 %%%%%%%%%%%%%%%%%

   fid_1=0;
   while fid_1<1                     %% Open file for reading
     [fid_1, message]=fopen(['st',st_str,'-',date_str,'-',st_dir,'-
4.dat'],'r');
     if (fopen(fid_1) == -1)
        disp(message);
     end
   end
   [A3,Count]=fscanf(fid_1,'%f %f %f %f',[N_state,N_dat]);
   A3=A3';
   fclose(fid_1);

%%%%%%%%%%%%%%% Lane 5 %%%%%%%%%%%%%%%%%

   fid_1=0;
   while fid_1<1                     %% Open file for reading
     [fid_1, message]=fopen(['st',st_str,'-',date_str,'-',st_dir,'-
5.dat'],'r');
     if (fopen(fid_1) == -1)
        disp(message);
     end
   end
   [A4,Count]=fscanf(fid_1,'%f %f %f %f',[N_state,N_dat]);
   A4=A4';
   fclose(fid_1);

%%%%%%%%%%%%%%% End %%%%%%%%%%%%%%%%%

%%% Convert the time index from float type to integer type. %%

A0(:,1)= round(A0(:,1));
A1(:,1)= round(A1(:,1));
A2(:,1)= round(A2(:,1));
A3(:,1)= round(A3(:,1));
A4(:,1)= round(A4(:,1));
```

```
A_Mean = zeros(Total_SampleTime,4);     %% value for each individual
second

Time_Index=zeros(No_Interval, 1);

Speed = zeros(No_Interval, 1);          %% aggregated value for each
sampling time interval
Occupancy = zeros(No_Interval, 1);
Flow = zeros(No_Interval, 1);
Density = zeros(No_Interval, 1);
Probe_Density = zeros(No_Interval, 1);  %% estimated density from VII
probe vehicles

Speed_Lane = zeros(No_Interval, 5);     %% aggregated value for each
individual lane
Flow_Lane = zeros(No_Interval, 5);      %% aggregated value for each
individual lane
Density_Lane = zeros(No_Interval, 5);   %% aggregated value for each
individual lane

%% Initialize the matrix
% Data_Lane_i stored the sorted data for the next step to calculate the
mean
% for lane group level at the specified time index

Data_Lane_1 = zeros(Total_SampleTime,4); % second_index, occupancy,
speed, flow
Data_Lane_2 = zeros(Total_SampleTime,4);
Data_Lane_3 = zeros(Total_SampleTime,4);
Data_Lane_4 = zeros(Total_SampleTime,4);
Data_Lane_5 = zeros(Total_SampleTime,4);

for i=1:Total_SampleTime
    Data_Lane_1(i,1)=i;
    Data_Lane_1(i,2)=-1;
    Data_Lane_1(i,3)=-1;
    Data_Lane_2(i,1)=i;
    Data_Lane_2(i,2)=-1;
    Data_Lane_2(i,3)=-1;
    Data_Lane_3(i,1)=i;
    Data_Lane_3(i,2)=-1;
    Data_Lane_3(i,3)=-1;
    Data_Lane_4(i,1)=i;
    Data_Lane_4(i,2)=-1;
    Data_Lane_4(i,3)=-1;
    Data_Lane_5(i,1)=i;
    Data_Lane_5(i,2)=-1;
    Data_Lane_5(i,3)=-1;

    A_Mean(i,1)=i;      % time index
    A_Mean(i,2)=-1;     % occupancy
    A_Mean(i,3)=-1;     % speed
    A_Mean(i,4)=0;      % vehicle count for one interval

    A_Mean(i,5)=0;      % accumulative vehicle count
```

```
    A_Mean(i,6)=0;      % accumulative vehicle count for lane 1
    A_Mean(i,7)=0;      % accumulative vehicle count for lane 2
    A_Mean(i,8)=0;      % accumulative vehicle count for lane 3
    A_Mean(i,9)=0;      % accumulative vehicle count for lane 4
    A_Mean(i,10)=0;     % accumulative vehicle count for lane 5

end

for i=1:size(A0)
    Data_Lane_1(A0(i,1)+1,2)=A0(i,2);
    Data_Lane_1(A0(i,1)+1,3)=A0(i,3);
    Data_Lane_1(A0(i,1)+1,4)=1;
end

for i=1:size(A1)
    Data_Lane_2(A1(i,1)+1,2)=A1(i,2);
    Data_Lane_2(A1(i,1)+1,3)=A1(i,3);
    Data_Lane_2(A1(i,1)+1,4)=1;
end

for i=1:size(A2)
    Data_Lane_3(A2(i,1)+1,2)=A2(i,2);
    Data_Lane_3(A2(i,1)+1,3)=A2(i,3);
    Data_Lane_3(A2(i,1)+1,4)=1;
end

for i=1:size(A3)
    Data_Lane_4(A3(i,1)+1,2)=A3(i,2);
    Data_Lane_4(A3(i,1)+1,3)=A3(i,3);
    Data_Lane_4(A3(i,1)+1,4)=1;
end

for i=1:size(A4)
    Data_Lane_5(A4(i,1)+1,2)=A4(i,2);
    Data_Lane_5(A4(i,1)+1,3)=A4(i,3);
    Data_Lane_5(A4(i,1)+1,4)=1;
end

%%%%%%%% to calculate the average vehicle count %%%%%%%
%%%% No_Interval = Total_SampleTime/DeltaT_SampleTime; % 5400 / 60 =
90, one hour and half

Sum_Speed = 0;
Sum_Occupancy = 0;
Vehicle_Count=0;
Vehicle_Count_1=0;
Vehicle_Count_2=0;
Vehicle_Count_3=0;
Vehicle_Count_4=0;
Vehicle_Count_5=0;

Interval_Vehicle_Count = 0;          % related to volume & density
Interval_Vehicle_Count_1=0;
Interval_Vehicle_Count_2=0;
Interval_Vehicle_Count_3=0;
```

```
Interval_Vehicle_Count_4=0;
Interval_Vehicle_Count_5=0;
Accumu_Vehicle_Count = 0;           % related to accumulative vechile
count

for j=1:No_Interval
    for i = 1:DeltaT_SampleTime
        if Data_Lane_1((j-1)*DeltaT_SampleTime+i,2)~=-1
            Flow_Lane(j,1)=Flow_Lane(j,1)+1;    % accutally it is
vehicle count now;
            Speed_Lane(j,1)=Speed_Lane(j,1)+ Data_Lane_1((j-
1)*DeltaT_SampleTime+i,3);  % actually it is speed sum now
            Vehicle_Count= Vehicle_Count+1;
            Vehicle_Count_1=Vehicle_Count_1+1;
            Interval_Vehicle_Count= Interval_Vehicle_Count+1;
            Sum_Occupancy = Sum_Occupancy + Data_Lane_1((j-
1)*DeltaT_SampleTime+i,2);
            Sum_Speed = Sum_Speed + Data_Lane_1((j-
1)*DeltaT_SampleTime+i,3);
        end
        if Data_Lane_2((j-1)*DeltaT_SampleTime+i,2)~=-1
            Flow_Lane(j,2)=Flow_Lane(j,2)+1;    % accutally it is
vehicle count now;
            Speed_Lane(j,2)=Speed_Lane(j,2)+ Data_Lane_2((j-
1)*DeltaT_SampleTime+i,3);  % actually it is speed sum now
            Vehicle_Count= Vehicle_Count+1;
            Vehicle_Count_2=Vehicle_Count_2+1;
            Interval_Vehicle_Count= Interval_Vehicle_Count+1;
            Sum_Occupancy = Sum_Occupancy + Data_Lane_2((j-
1)*DeltaT_SampleTime+i,2);
            Sum_Speed = Sum_Speed + Data_Lane_2((j-
1)*DeltaT_SampleTime+i,3);
        end
        if Data_Lane_3((j-1)*DeltaT_SampleTime+i,2)~=-1
            Flow_Lane(j,3)=Flow_Lane(j,3)+1;    % accutally it is
vehicle count now;
            Speed_Lane(j,3)=Speed_Lane(j,3)+ Data_Lane_3((j-
1)*DeltaT_SampleTime+i,3);  % actually it is speed sum now
            Vehicle_Count= Vehicle_Count+1;
            Vehicle_Count_3=Vehicle_Count_3+1;
            Interval_Vehicle_Count= Interval_Vehicle_Count+1;
            Sum_Occupancy = Sum_Occupancy + Data_Lane_3((j-
1)*DeltaT_SampleTime+i,2);
            Sum_Speed = Sum_Speed + Data_Lane_3((j-
1)*DeltaT_SampleTime+i,3);
        end
        if Data_Lane_4((j-1)*DeltaT_SampleTime+i,2)~=-1
            Flow_Lane(j,4)=Flow_Lane(j,4)+1;    % accutally it is
vehicle count now;
            Speed_Lane(j,4)=Speed_Lane(j,4)+ Data_Lane_4((j-
1)*DeltaT_SampleTime+i,3);  % actually it is speed sum now
            Vehicle_Count= Vehicle_Count+1;
            Vehicle_Count_4=Vehicle_Count_4+1;
            Interval_Vehicle_Count= Interval_Vehicle_Count+1;
            Sum_Occupancy = Sum_Occupancy + Data_Lane_4((j-
1)*DeltaT_SampleTime+i,2);
```

```
            Sum_Speed = Sum_Speed + Data_Lane_4((j-
1)*DeltaT_SampleTime+i,3);
        end
        if Data_Lane_5((j-1)*DeltaT_SampleTime+i,2)~=-1
            Flow_Lane(j,5)=Flow_Lane(j,5)+1;    % accutally it is
vehicle count now;
            Speed_Lane(j,5)=Speed_Lane(j,5)+ Data_Lane_5((j-
1)*DeltaT_SampleTime+i,3);  % actually it is speed sum now
            Vehicle_Count= Vehicle_Count+1;
            Vehicle_Count_5=Vehicle_Count_5+1;
            Interval_Vehicle_Count= Interval_Vehicle_Count+1;
            Sum_Occupancy = Sum_Occupancy + Data_Lane_5((j-
1)*DeltaT_SampleTime+i,2);
            Sum_Speed = Sum_Speed + Data_Lane_5((j-
1)*DeltaT_SampleTime+i,3);
        end

        Accumu_Vehicle_Count=Accumu_Vehicle_Count+Vehicle_Count;
        A_Mean((j-1)*DeltaT_SampleTime+i,5)=Accumu_Vehicle_Count;
        A_Mean((j-1)*DeltaT_SampleTime+i,6)=Vehicle_Count_1;
        A_Mean((j-1)*DeltaT_SampleTime+i,7)=Vehicle_Count_2;
        A_Mean((j-1)*DeltaT_SampleTime+i,8)=Vehicle_Count_3;
        A_Mean((j-1)*DeltaT_SampleTime+i,9)=Vehicle_Count_4;
        A_Mean((j-1)*DeltaT_SampleTime+i,10)=Vehicle_Count_5;

        Vehicle_Count = 0;
    end
        for m=1:5
            Speed_Lane(j,m)=Speed_Lane(j,m)/Flow_Lane(j,m);
            Flow_Lane(j,m)=Flow_Lane(j,m)*3600/DeltaT_SampleTime;
        end

    Time_Index(j,1)=j;
    A_Mean(j,4)=Interval_Vehicle_Count;
    Speed(j,1)=Sum_Speed/(Interval_Vehicle_Count+0.00000000000001);
    Occupancy(j,1)=Sum_Occupancy/Total_SampleTime;
    Flow(j,1)=Interval_Vehicle_Count*3600/(5*DeltaT_SampleTime);
    Density(j,1)=Flow(j,1)/Speed(j,1);
    Sum_Speed = 0;
    Interval_Vehicle_Count = 0;
end

%%%%%%%% to calculate the average vechile count %%%%%%%%

%%%%%%%%%%%%%% Data imputation %%%%%%%%%%%%%%%%%%%%%%%%%

for j=2:No_Interval
    if Speed(j,1)==0
        Speed(j,1)=Speed(j-1,1);
    end
    Density(j,1)=Flow(j,1)/Speed(j,1);

    for m=1:5
        if Speed_Lane(j,m)==0
        Speed_Lane(j,m)=Speed_Lane(j-1,m);
```

```
        end

Density_Lane(j,m)=Flow_Lane(j,m)/(Speed_Lane(j,m)+0.0000000001);
    end
end

for i=2:Total_SampleTime
    if Data_Lane_1(i,2)==-1
        Data_Lane_1(i,2)=Data_Lane_1(i-1,2);
    end
    if Data_Lane_1(i,3)==-1
        Data_Lane_1(i,3)=Data_Lane_1(i-1,3);
    end

    if Data_Lane_2(i,2)==-1
        Data_Lane_2(i,2)=Data_Lane_2(i-1,2);
    end
    if Data_Lane_2(i,3)==-1
        Data_Lane_2(i,3)=Data_Lane_2(i-1,3);
    end

    if Data_Lane_3(i,2)==-1
        Data_Lane_3(i,2)=Data_Lane_3(i-1,2);
    end
    if Data_Lane_3(i,3)==-1
        Data_Lane_3(i,3)=Data_Lane_3(i-1,3);
    end

    if Data_Lane_4(i,2)==-1
        Data_Lane_3(i,2)=Data_Lane_4(i-1,2);
    end
    if Data_Lane_4(i,3)==-1
        Data_Lane_3(i,3)=Data_Lane_4(i-1,3);
    end

    if Data_Lane_5(i,2)==-1
        Data_Lane_5(i,2)=Data_Lane_5(i-1,2);
    end
    if Data_Lane_5(i,3)==-1
        Data_Lane_5(i,3)=Data_Lane_5(i-1,3);
    end

end
%%%%%%%%%%%%%%% Data imputation %%%%%%%%%%%%%%%%%%%%%%%%

for i=1:Total_SampleTime

    sum_Occupancy=0;
    sample_Occupancy = 0;

    sum_Speed=0;
    sample_Speed = 0;

    %%%%%%%% to calculate the average occupancy %%%%%%%%
```

```
    if Data_Lane_1(i,2)~=-1
        sum_Occupancy = sum_Occupancy + Data_Lane_1(i,2);
        sample_Occupancy=sample_Occupancy+1;
    end
    if Data_Lane_2(i,2)~=-1
        sum_Occupancy = sum_Occupancy + Data_Lane_2(i,2);
        sample_Occupancy=sample_Occupancy+1;
    end
    if Data_Lane_3(i,2)~=-1
        sum_Occupancy = sum_Occupancy + Data_Lane_3(i,2);
        sample_Occupancy=sample_Occupancy+1;
    end
    if Data_Lane_4(i,2)~=-1
        sum_Occupancy = sum_Occupancy + Data_Lane_4(i,2);
        sample_Occupancy=sample_Occupancy+1;
    end
    if Data_Lane_5(i,2)~=-1
        sum_Occupancy = sum_Occupancy + Data_Lane_5(i,2);
        sample_Occupancy=sample_Occupancy+1;
    end

    A_Mean(i,2)=sum_Occupancy/(sample_Occupancy+0.000000000001);

    %%%%%%%% to calculate the average speed %%%%%%%

    if Data_Lane_1(i,3)~=-1
        sum_Speed = sum_Speed + Data_Lane_1(i,3);
        sample_Speed=sample_Speed+1;
    end
    if Data_Lane_2(i,3)~=-1
        sum_Speed = sum_Speed + Data_Lane_2(i,3);
        sample_Speed=sample_Speed+1;
    end
    if Data_Lane_3(i,3)~=-1
        sum_Speed = sum_Speed + Data_Lane_3(i,3);
        sample_Speed=sample_Speed+1;
    end
    if Data_Lane_4(i,3)~=-1
        sum_Speed = sum_Speed + Data_Lane_4(i,3);
        sample_Speed=sample_Speed+1;
    end
    if Data_Lane_5(i,3)~=-1
        sum_Speed = sum_Speed + Data_Lane_5(i,3);
        sample_Speed=sample_Speed+1;
    end

    A_Mean(i,3)=sum_Speed/(sample_Speed++0.000000000001);
%    A_Mean(i,4)=sample_Speed;
end

    A_Mean(:,2)=butt_flt(A_Mean(:,2),15);
    A_Mean(:,3)=butt_flt(A_Mean(:,3),15);

%%%%%%%%%%%%%%% Save the mean data into a .dat file %%%%%%%%%%%%%%%%%
```

```
    fid_mean=0;
    while fid_mean<1                         %% Open file for writing
       [fid_mean, message]=fopen(['st',st_str,'-',st_dir,'-
mean.dat'],'w+');
      if (fopen(fid_mean) == -1)
         disp(message);
      end
    end

    for i=1:size(A_Mean)
        fprintf(fid_mean, '%d\t %f\t %f\t %d\t %d\t %d\t %d\t %d\t %d\t
%d\n',A_Mean(i,1),A_Mean(i,2),A_Mean(i,3),A_Mean(i,4),A_Mean(i,5),A_Mea
n(i,6),A_Mean(i,7),A_Mean(i,8),A_Mean(i,9),A_Mean(i,10));
    end

    fclose(fid_mean);     %% Close written file

%%%%%% Plot the figures %%%%%%%%%%%%

    figure(1);

    subplot(3,1,1);
    axis([1,Total_SampleTime,0,1]);
    hold on;
    plot(A_Mean(:,1), A_Mean(:,2));
    set(gca,'YTick',0:0.2:1);
    set(gca,'XTick',0:900:Total_SampleTime);

set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
    grid;
    hold off;
    ylabel(['Mean-Occupancy @ Station -',st_str]);

    subplot(3,1,2);
    axis([1,Total_SampleTime,0,100]);
    hold on;
    plot(A_Mean(:,1), A_Mean(:,3));
    set(gca,'YTick',0:20:100);
    set(gca,'XTick',0:900:Total_SampleTime);

set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
    grid;
    hold off;
    ylabel(['Mean-Speed @ Station-',st_str]);

    subplot(3,1,3);
    axis([1,Total_SampleTime,0,2400]);
    hold on;
    plot(A_Mean(:,1)*DeltaT_SampleTime, A_Mean(:,4)*12);
    set(gca,'YTick',0:300:2400);
    set(gca,'XTick',0:900:Total_SampleTime);
```

```
set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
    grid;
    hold off;
    ylabel(['Mean-Flow @ Station-',st_str]);

%%%%%% Plot the figures %%%%%%%%%%%%

    figure(2);

    subplot(3,1,1);
    axis([1,No_Interval,0,100]);
    hold on;
    plot(Density(:,1));
    set(gca,'YTick',0:10:100);
    set(gca,'XTick',0:(No_Interval/6):No_Interval);

set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
    grid;
    hold off;
%   ylabel(['Density @ Station -',st_str]);
    ylabel('Density (veh/mile)');
    xlabel('Density between Station 6 and Station 7');

    subplot(3,1,2);
    axis([1,No_Interval,0,100]);
    hold on;
    plot(Speed(:,1));
    set(gca,'YTick',0:20:100);
    set(gca,'XTick',0:(No_Interval/6):No_Interval);

set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
    grid;
    hold off;
    ylabel(['Speed (mph)']);
%   ylabel(['Speed @ Station-',st_str]);
    xlabel('Speed between Station 6 and Station 7');

    subplot(3,1,3);
    axis([1,No_Interval,0,2400]);
    hold on;
    plot(Flow(:,1));
    set(gca,'YTick',0:300:2400);
    set(gca,'XTick',0:(No_Interval/6):No_Interval);

set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
    grid;
    hold off;
    ylabel('Flow (veh/h/l)');
    xlabel('Flow at Station 6');

%%%%%% Plot the figures for density comparison %%%%%%%%%%%%
```

```
   fid_2=0;
   while fid_2<1                      %% Open file for reading
     [fid_2, message]=fopen('ComDensity_1.txt','r');
     if (fopen(fid_2) == -1)
        disp(message);
     end
   end
   [Com_Density_1,Count]=fscanf(fid_2,'%d %d',[2,5]);
   Com_Density_1=Com_Density_1';
   fclose(fid_2);

fid_2=0;
   while fid_2<1                      %% Open file for reading
     [fid_2, message]=fopen('ComDensity_2.txt','r');
     if (fopen(fid_2) == -1)
        disp(message);
     end
   end
   [Com_Density_2,Count]=fscanf(fid_2,'%d %d',[2,5]);
   Com_Density_2=Com_Density_2';
   fclose(fid_2);

   Probe_Lane_Density = zeros(2,2,5); %% (density,time index, lane
number)

   Probe_Lane_Density(1,1,5)=Com_Density_1(1,1);
   Probe_Lane_Density(1,2,5)=Com_Density_1(1,2);
   Probe_Lane_Density(2,1,5)=Com_Density_2(1,1);
   Probe_Lane_Density(2,2,5)=Com_Density_2(1,2);

   Probe_Lane_Density(1,1,4)=Com_Density_1(2,1);
   Probe_Lane_Density(1,2,4)=Com_Density_1(2,2);
   Probe_Lane_Density(2,1,4)=Com_Density_2(2,1);
   Probe_Lane_Density(2,2,4)=Com_Density_2(2,2);

   Probe_Lane_Density(1,1,3)=Com_Density_1(3,1);
   Probe_Lane_Density(1,2,3)=Com_Density_1(3,2);
   Probe_Lane_Density(2,1,3)=Com_Density_2(3,1);
   Probe_Lane_Density(2,2,3)=Com_Density_2(3,2);

   Probe_Lane_Density(1,1,2)=Com_Density_1(4,1);
   Probe_Lane_Density(1,2,2)=Com_Density_1(4,2);
   Probe_Lane_Density(2,1,2)=Com_Density_2(4,1);
   Probe_Lane_Density(2,2,2)=Com_Density_2(4,2);

   Probe_Lane_Density(1,1,1)=Com_Density_1(5,1);
   Probe_Lane_Density(1,2,1)=Com_Density_1(5,2);
   Probe_Lane_Density(2,1,1)=Com_Density_2(5,1);
   Probe_Lane_Density(2,2,1)=Com_Density_2(5,2);

   figure(3);

    axis([1,5400,0,100]);
    hold on;
```

```
plot(Com_Density_1(:,2),Com_Density_1(:,1),'b*',Time_Index(:,1)*DeltaT_
SampleTime,Density(:,1),'g-');
    set(gca,'YTick',0:10:100);
    set(gca,'XTick',0:900:5400);

set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
    legend('Density from Loop and Probe Vehicle','Density from Loop
Data');
    grid;
    hold off;
    xlabel('Estimated Density between Station 6 and 7');
    ylabel(['Density (veh/mile)']);

    fid_2=0;
   while fid_2<1                          %% Open file for reading
     [fid_2, message]=fopen('Probe_Vehicle.txt','r');
     if (fopen(fid_2) == -1)
        disp(message);
     end
   end
   N_Trips=10;
   Length = 334;
   Lane_Number = 5;
   Probe_Station_Density=zeros(1, N_Trips);
   [T_Probe,Count]=fscanf(fid_2,'%d %d',[2,N_Trips]);
   T_Probe=T_Probe';  % T_Probe(:,1) arrival time for upstream station;
T_Probe(:,2) arrival time for downstream station
   fclose(fid_2);
    for m=1:N_Trips
        % accumulative vehicle count change
        Probe_Station_Density(m)=round((A_Mean(T_Probe(m,2),5)-
A_Mean(T_Probe(m,1),5))*1600/(334*Lane_Number));
    end

    for j=1:No_Interval
        Probe_Density(j)=-1;
        Smpling_Probe_No = 0;
        Sum_Probe_Density = 0;
        for m=1:N_Trips
            if (T_Probe(m,2)>(j-
1)*DeltaT_SampleTime)&(T_Probe(m,2)<=j*DeltaT_SampleTime)
                Smpling_Probe_No=Smpling_Probe_No+1;

Sum_Probe_Density=Sum_Probe_Density+Probe_Station_Density(m);
            end
        end
        if Smpling_Probe_No>0
           Probe_Density(j)=Sum_Probe_Density/Smpling_Probe_No;
        end
    end

    figure(4);
    axis([1,5400,0,100]);
    hold on;
```

```
plot(Time_Index(:,1)*DeltaT_SampleTime,Probe_Density(:,1),'b*',Time_Ind
ex(:,1)*DeltaT_SampleTime,Density(:,1),'g-');
    set(gca,'YTick',0:10:100);
    set(gca,'XTick',0:900:5400);

set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
    legend('Density from Loop and Probe Vehicle','Density from Loop
Data');
    grid;
    hold off;
    xlabel('Estimated Density between Station 6 and 7');
    ylabel(['Density (veh/mile)']);

    figure(5);
    for m=1:5
    %Density_Lane(:,m)=butt_flt(Density_Lane(:,m),800);
    subplot(5,1,m);
    %figure(3+m);
    axis([1,5400,0,100]);
    hold on;

plot(Probe_Lane_Density(:,2,m),Probe_Lane_Density(:,1,m),'r*',Time_Inde
x(:,1)*DeltaT_SampleTime,Density_Lane(:,m),'b-');
    set(gca,'YTick',0:20:100);
    set(gca,'XTick',0:900:5400);

set(gca,'XTickLabel',{'9:00','9:15','9:30','9:45','10:00','10:15','10:3
0'});
 %   legend('Density from Loop and Probe Vehicle','Density from Loop
Data');
    grid;
    hold off;
    xlabel(['Estimated Density for Lane ',int2str(m), ' between Station
6 and 7']);
    ylabel(['Density (veh/mile)']);
    end
```