

**REAL-TIME PREDICTION OF QUEUES AT SIGNALIZED
INTERSECTIONS TO SUPPORT ECO-DRIVING
APPLICATIONS**

Final Report



TranLIVE

Mecit Cetin, Ozhan Unal

October 2014

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Real-time prediction of queues at signalized intersections to support eco-driving applications		5. Report Date October 2014	
		6. Performing Organization Code KLG900-SB-001	
7. Author(s) Cetin, Mecit; Unal, Ozhan		8. Performing Organization Report No. N14-18	
9. Performing Organization Name and Address Transportation Research Institute (TRI) Old Dominion University 4111 Monarch Way, Suite 204 Norfolk, VA 23506 Source Organization Name and Address TranLIVE/NIATT University of Idaho 875 Perimeter Dr. MS0901 Moscow, ID 83844-0901		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. DTRT12GUTC17	
12. Sponsoring Agency Name and Address US Department of Transportation Research and Special Programs Administration 400 7th Street SW Washington, DC 20509-0001		13. Type of Report and Period Covered Final Report: August 2012– September 2014	
		14. Sponsoring Agency Code USDOT/RSPA/DIR-1	
15. Supplementary Notes:			
16. Abstract The overall objective of this research is to develop models for predicting queue lengths at signalized intersections based on the data from probe vehicles. The time and space coordinates of the probe vehicles going through signalized intersections are utilized to predict the back of the queue profile. For a single intersection, prediction models are developed where both over-saturated and under-saturated conditions are considered. The shockwave theory (i.e., the Lighthill-Whitham-Richards theory) is used to estimate the evolution of the back of the queue over time and space from the event data generated when probe vehicles join the back of the queue. An analytical formulation is developed for determining the critical points required to create the time-space diagrams that characterize queue dynamics. These critical points are used to estimate the queue lengths. The formulation is tested on the data obtained from traffic simulation software VISSIM. It was found that the shockwave-based formulation is effective in estimating queue dynamics at signalized intersections for -- and over-saturated conditions even with a relatively low percentage of probes (e.g., 10-20%) in the system. For example, under over-saturated conditions simulated, the error is less than $\pm 10\%$ in more 90% of the cycles when the market penetration of probe vehicles is 15%.			
17. Key Words Probe vehicles, queue lengths, traffic signals		18. Distribution Statement Unrestricted; Document is available to the public through the National Technical Information Service; Springfield, VT.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 24	22. Price ...

TABLE OF CONTENTS

EXECUTIVE SUMMARY 1

INTRODUCTION 3

PROBLEM DEFINITION AND SETTING..... 6

METHODOLOGY 8

APPLICATION TO SIMULATION DATA 14

DISCUSSION OF RESULTS 20

CONCLUSIONS..... 21

ACKNOWLEDGEMENTS 21

REFERENCES 22

APPENDIX..... 24

FIGURES

Figure 1: Vehicle trajectories and shockwave diagram (both under-saturated and over-saturated conditions exist)..... 8

Figure 2: Shockwave lines at traffic signal under over-saturated conditions. 9

Figure 3: A closer look to time-space diagram during green and red phases. 11

Figure 4: Estimated back of queue and shockwave lines with all vehicles (30% probe). 15

Figure 5: Estimated back of queue and shockwave lines with all vehicles (20% probe). 15

Figure 6: Estimated back of queue and shockwave lines with all vehicles (10% probe). 16

Figure 7: Back of queue estimated with all vehicles and 10% probes (input volume 1150 vph). 17

Figure 8: Ten different back of queue profiles estimated with randomly selected 5% probes. 18

TABLES

Table 1: Summary Statistics for Queue Length Estimation at Different Probe Levels 17

Table 2: Summary Statistics for Back of Queue Prediction at Different Probe Levels (input volume 1150 vph) 19

EXECUTIVE SUMMARY

For signalized intersections, queue length is one of the most important performance measures. Knowing the evolution of queue lengths over time and space allows quantifying system performance and improving signal operations, and supports eco-driving and eco-signals applications. Previously, researchers estimated the average queue length at traffic signals based on loop detector data and signal timing information. This study is focused on estimating queue lengths using the information provided by probe vehicles. As more vehicles become equipped with location tracking and communications systems, they will provide new opportunities to observe the transportation networks more effectively. The United States Department of Transportation's Connected Vehicle program is a significant effort to make the vision of seamless vehicle-to-vehicle and vehicle-to-infrastructure communication a reality. Within this vision, vehicles will be aware of their own locations in the transportation systems and will exchange useful information with other vehicles as well as with the infrastructure. By tracking the positions of these "probe" vehicles along roadway segments a wealth of information is generated to precisely characterize traffic flow dynamics. This in turn allows the system operators to improve system efficiency by taking relevant control actions (e.g., retiming traffic signals, responding faster to incidents).

In this study, the time and space coordinates of those probe vehicles going through signalized intersections are utilized to predict the back of the queue profile. For a single intersection, prediction models are developed where both over-saturated and under-saturated conditions are considered. The shockwave theory (i.e., the Lighthill-Whitham-Richards theory) is used to estimate the evolution of the back of the queue over time and space from the event data generated when probe vehicles join the back of the queue. An analytical formulation was developed for determining the critical points required to create time-space diagrams to characterize the queue dynamics. These critical points are used to estimate the queue lengths. The formulation was tested on the data obtained from traffic simulation software VISSIM. It was found that the shockwave-based formulation is effective in estimating queue dynamics at signalized intersections for under- and over-saturated conditions even with a relatively low percentage of probes (e.g., 10-20%) in the system. For example, under over-saturated conditions

simulated, the error is less than $\pm 10\%$ in more 90% of the cycles when the market penetration of probe vehicles is 15%.

INTRODUCTION

The overall objective of this research is to develop models for predicting queue lengths at signalized intersections based on the data from probe vehicles. As more vehicles become equipped with location tracking and communications systems, they will provide new opportunities to observe the transportation networks more effectively. The United States Department of Transportation's Connected Vehicle program is a significant effort to make the vision of seamless vehicle-to-vehicle and vehicle-to-infrastructure communication a reality. Within this vision vehicles will be aware of their own locations in the transportation systems and will exchange useful information with other vehicles as well as with the infrastructure. By tracking the positions of these "probe" vehicles along roadway segments a wealth of information is generated to precisely characterize traffic flow dynamics. This in turn allows the system operators to improve system efficiency by taking relevant control actions (e.g., retiming traffic signals, responding faster to incidents). In addition, knowing the evolution of (downstream) queues over time allows equipped vehicles to adjust their speeds to avoid stopping at the signals (if possible) or to minimize the changes to their speed profiles. This will result in less gas consumption and less tailpipe emissions. Such systems or strategies are referred to as "eco-traffic signal systems" or "eco-driving."

Recently, various studies have investigated the use of probe vehicle data for queue length estimation (1-7). In this project, probe vehicle data are also used in order to study queuing phenomena observed at traffic signals. More specifically, time and space data gathered from probe vehicles are used to estimate the back of the queue over time. Shockwave theory is used to estimate the evolution of the back of the queue over both time and space from the data when probe vehicles join the back of the queue (3, 8). Cumulative curve method is another approach for studying queue dynamics (9). In this method, all vehicles at specific locations need to be counted which is not possible when market penetration of probe vehicles is less than 100%. Moreover, this approach does not represent the spatial extent of the queue explicitly. Although there are some studies that show that the cumulative count curves can be used to determine the spatial extent of the queue (9, 10), when the arrival rate is time dependent, it is easier to construct the back of the queue by using shockwave theory.

The kinematic wave model which is known as Lighthill-Whitham-Richards (LWR) model in the literature is used in this study to describe traffic flow dynamics (11-13). LWR model is based on the principle of conservation of vehicles and a fundamental diagram that relates flow to density. This model can be used to analyze important traffic flow phenomena (e.g., shockwaves). In various studies, the original LWR model has been extended to describe further complexities (14-17).

Liu et al. use the LWR shockwave theory with high-resolution traffic signal data to distinguish different traffic states at the intersection so that queue length is estimated under congested conditions (14). Ban et al. propose methods to estimate real-time queue lengths at signalized intersections using sample travel times from mobile traffic sensors between some predefined virtual points before and after the signalized intersections (1). They identify the critical points when the queue is maximum, minimum or cleared within a cycle by using the delay patterns. There are additional studies that use LWR theory in estimating the queue lengths (e.g.,(18)).

Statistical methods are also used in estimating queue lengths at signalized intersections. Comert and Cetin proposed analytical models for errors in queue length estimation using probe vehicle data (6, 7). Some studies assumed discrete arrivals and integer cycle lengths, and Markov chain or similar statistical analysis techniques were applied to estimate the mean or distribution of queue lengths. Some studies estimated the average queue length of a fixed-time signal by assuming traffic flow and signal timing parameters are continuous variables (19). However, these models are point queue models and do not represent the space or distance explicitly.

In a recent study, Cetin and Rakha (20) present methods to estimate the total fuel consumption and CO₂ emissions at a signalized intersection from the probe vehicle data. Traffic flow through an intersection is simulated to generate vehicle trajectories under both congested and uncongested conditions. By using the Virginia Tech Comprehensive Power-Based Fuel Consumption Model (21), the total fuel consumed by each vehicle is determined for a given trajectory. Several alternative methods are presented to estimate the total fuel consumption from the sample data provided by the probe vehicles. Their results show that a simple extrapolation of the fuel consumed by probes to the rest of the traffic does not yield very accurate results. A more accurate solution is obtained by capitalizing on the probe trajectories to construct trajectories for

the non-probe vehicles. For the simulated conditions, it is demonstrated that the total fuel consumption can be estimated with a reasonable accuracy at relatively low probe-vehicle market-penetration levels. It is further demonstrated that if a proper “average vehicle” is specified for estimating the total fuel consumption level, then knowing the make & model of individual probe vehicles does not enhance the estimation accuracy. The models presented by Cetin and Rakha (20) can support data collection systems needed for eco-signals and other similar applications to improve fuel consumption and to reduce CO₂ emissions which are directly proportional to the fuel usage.

The methodology proposed in this study differs from the previous studies. The queue length is estimated by using the time-space coordinates of the probe vehicles when they join the back of the queue in each cycle. For over-saturated conditions, data only from the first and last probe vehicles are used to estimate the queue length while only the last probe vehicle data is used for under-saturated conditions. The methodology does not make any assumptions about the percentage of the probe vehicles. Moreover, it does not require probe observations in each cycle, in contrast to the minimum requirement of two probe observations per cycle as indicated by Ban et al.(1).

Cetin proposed a methodology to estimate the queue lengths for over-saturated conditions at a signalized intersection (3). This study is an extension of the aforementioned study to account for number of cycles which can be either in over-saturated or under-saturated conditions.

PROBLEM DEFINITION AND SETTING

To develop the formulation for queue length estimation, some simplifying assumptions are made. Figure 1 shows a sample time–space diagram with vehicle trajectories in which the cycles are either in under-saturated or over-saturated conditions. Based on the LWR theory, shockwave lines between different traffic states are also shown in the figure. Critical points indicated by R_n and Q_n , characterize the dynamics of the back of the queue, where n denotes the cycle number. The assumptions under which the formulation is developed are described in the following subsections.

In this section the assumptions to derive the queue estimation model are presented.

- **Roadway geometry and signal timing:**
A single-lane road leading to a signalized intersection is considered. The traffic signal is assumed to operate on a fixed cycle with alternating green and red phases of equal durations. This should not be considered as a limitation of the work since the formulation can be easily adapted to varying cycle and phase lengths.
- **Probe data:**
Even though probe vehicles can technically collect data continuously (e.g., every second), such large dataset is not needed to estimate the queue dynamics. In this paper, only the location and time data when probe vehicles join the back of the queue are used. In other words, for each probe vehicle, a single data point is needed, which contains its location on the link and the time instant when it joins the back of the queue.
- **Probe vehicle population:**
There is no assumption on the percentage of the probe vehicles joining the back of the queue.
- **Traffic flow:**
It is assumed that the vehicles arrive randomly at the intersection with an unknown rate. The method presented here does not consider scenarios in which vehicle platoons form because of an upstream signal.

- **Shockwave speed:**

The shockwave speed is constant and the same for each cycle. Triangular fundamental diagram is used to find the shockwave speed. This will be explained in detail in the methodology section.

In VISSIM data, vehicles are assumed to be stopped when their speeds are less than 5 km/hr. For over-saturated conditions, speeds of the backward-moving shockwaves A_nR_n and B_nQ_n in Figure 1 are assumed to be known and equal. This assumption is based on the LWR theory. Based on the LWR theory, shockwave speed is computed by $\Delta q/\Delta k$ where

Δq = the difference in flows of the two traffic states separated by shockwave

Δk = the difference in densities

These differences will be the same for shockwaves A_nR_n and B_nQ_n in Figure 1.

Therefore, the shockwave speeds will also be the same.

- **Online Versus Offline Application**

The methodology proposed in this study can be applied for both offline and online applications. When applied online, the results obtained from probe vehicles arriving in the previous cycles can be used for predicting queue lengths. In this report, queues are estimated whenever a new probe vehicle joins the back of the queue.

- **Queue Conditions**

The formulations for over-saturated conditions (when there is a residual queue) and under-saturated conditions are different.

METHODOLOGY

Time and space coordinates from probe vehicles when they join back of the queue are used in the formulation. The critical points R_n and Q_n for over-saturated and Q_n for under-saturated conditions are estimated (R_n for under-saturated condition is already known since it corresponds to the start of red phase). Once these are estimated, the queue length can be computed since these points define the boundary points of the queues. In order to find these unknown points, formulations are derived based on LWR theory. In this section, the formulations for both under-saturated and over-saturated conditions are presented.

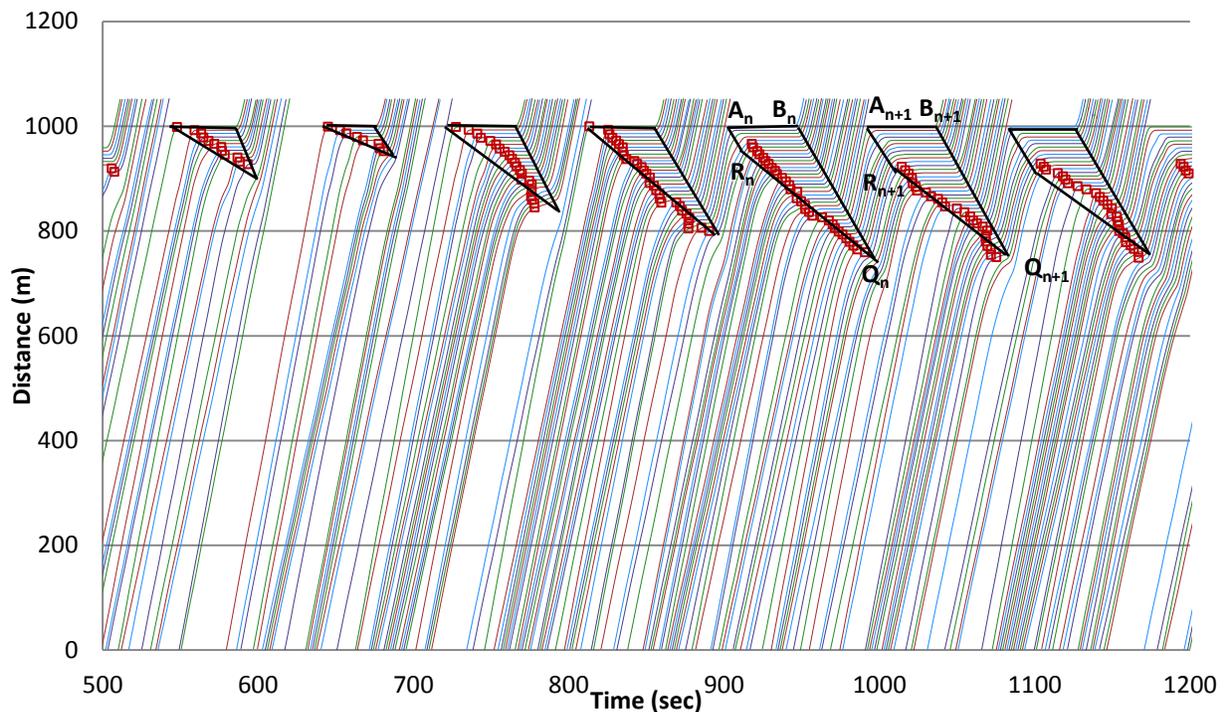


Figure 1: Vehicle trajectories and shockwave diagram (both under-saturated and over-saturated conditions exist).

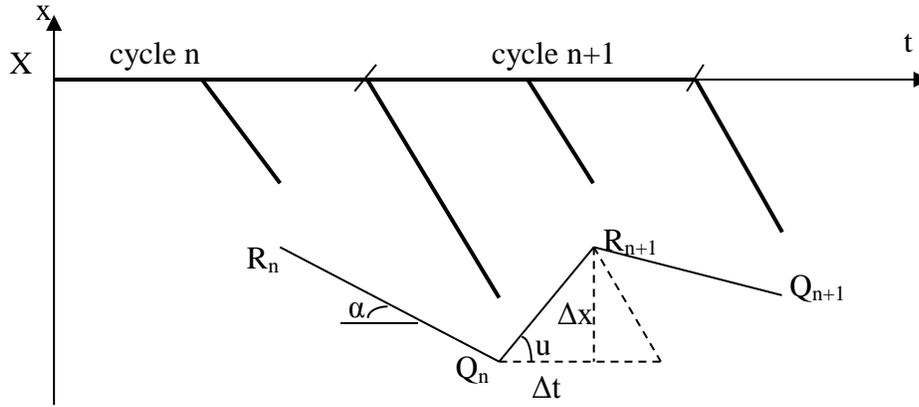


Figure 2: Shockwave lines at traffic signal under over-saturated conditions.

Shockwaves representing queue growth (i.e., all $R_n Q_n$ lines) are unknown for all cycles. The goal is to determine these shockwave speeds (or the slopes for the straight $R_n Q_n$ line segments) by using probe vehicle data. For over-saturated conditions, the first and the last probe vehicle observations (if any) are identified for each cycle. The last probe data in cycle n and the first probe vehicle data in cycle $n+k$ ($n+k$ is the next cycle with some probe vehicle observation) are used to find a constant shockwave speed (α in Figure 2). Since no information is available between these two points, the arrival rate is assumed to be constant. This assumption is realistic since there is no additional information to be used. After computing this constant speed α , the coordinates for all critical points for over-saturated cycles between these two probe observations are estimated. The formulation for over-saturated conditions is explained in more details in a previous study (3).

The formulas for over-saturated conditions (referring to Figure 2 below) are given as follows:

$$t_n^Q = \frac{(x_n^L - x_0) + wnC + \alpha t_n^L}{w + \alpha} \quad (1)$$

$$x_n^Q = x_0 + w(t_n^Q - nC) \quad (2)$$

$$\alpha = \frac{\Delta x - (x_{n+1}^F - x_n^L)}{(t_{n+1}^F - t_n^L) - \Delta t} \quad (3)$$

Where

x_n^Q and t_n^Q denote the space and time coordinates for point Q_n

X_0 = space coordinate of stop bar

w = shockwave speed

C = cycle length

x_n^L and t_n^L = the space and the time coordinates of the last probe vehicle in cycle n

x_{n+1}^L and t_{n+1}^L = the space and the time coordinates of the first probe vehicle in cycle $n+1$

α = the tangent of the angle shown in figure

Once x_n^Q and t_n^Q are found, X_{n+1}^R and t_{n+1}^R are calculated as

$$\Delta t = \frac{wG}{w-u} \quad (4)$$

$$\Delta x = u \frac{wG}{w-u} \quad (5)$$

$$X_{n+1}^R = X_n^Q + \Delta x \quad (6)$$

$$t_{n+1}^R = t_n^Q + \Delta t \quad (7)$$

where u is the shockwave speed shown in Figure 2 and G is the length of green phase.

Equations (1) through (7) are for the scenario in which there are probe vehicle observations in two consecutive cycles. A more general formulation is for the scenario in which there are probe vehicle observations in the n th and $(n+k)$ th cycles. Let n and $n+k$ be any two cycles for which probe vehicle observations are available. The slope α is calculated as;

$$\alpha = \frac{k\Delta x - (x_{n+k}^F - x_n^L)}{(t_{n+k}^F - t_n^L) - k\Delta t} \quad (8)$$

Once α is determined from Equation (8), coordinates of Q_n can be found by using Equations (1) and (2). R_{n+1} can be found by using equations (6) and (7). Finally x_{n+1}^Q and t_{n+1}^Q can be found as follows:

$$t_{n+1}^Q = \frac{\alpha t_{n+1}^R + x_{n+1}^R - x_0 + w(n+1)C}{w + \alpha} \tag{9}$$

$$x_{n+1}^Q = x_0 + w(t_{n+1}^Q - (n + 1)C) \tag{10}$$

Under-saturated Condition

To illustrate how deterministic queuing analysis can be employed for under-saturated condition, Figure 3 below illustrates the evolution of a queue of vehicles at an under-saturated signalized intersection where vehicles arrive at a uniform rate.

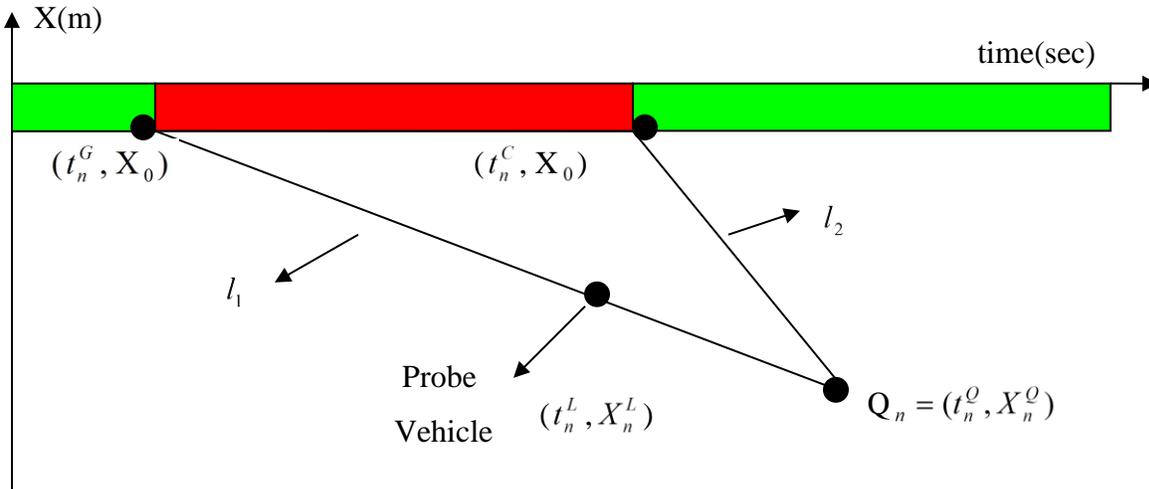


Figure 3: A closer look to time-space diagram during green and red phases.

For calculations, the triangle in Figure 3 is used, which shows the variables needed for the formulation. In order to estimate the critical points for under-saturated conditions, the intersection of two straight lines is found.

Q_n is the point where the maximum queue reaches. As mentioned before, the purpose of this study is to derive a formulation to find the coordinates of this point. In order to find the coordinates of Q_n , the equations for lines l_1 and l_2 are obtained. The coordinates for the end of the n^{th} cycle green phase, (t_n^G, X_0) in the figure, is known.

$$t_n^G = G + (n-1)C \quad n=1, 2, \dots \quad (11)$$

where G : Green time

C : Cycle length

Since the coordinates of last probe vehicle which is denoted by (t_n^L, X_n^L) on the figure, equation for l_1 can be written as;

$$X_n^Q - X_o = \left(\frac{X_n^L - X_0}{t_n^L - t_n^G} \right) (t_n^Q - t_n^G) \quad (12)$$

Where

$\left(\frac{X_n^L - X_0}{t_n^L - t_n^G} \right)$ is the slope of the l_1 and will be denoted by α

t_n^Q : The time coordinate of Q_n

X_n^Q : The space coordinate of Q_n

Equation of l_2 can be written as;

$$X_n^Q - X_o = w(t_n^Q - t_n^C) \quad (13)$$

where $t_n^C = nC \quad n = 1, 2, \dots$

w = shockwave speed

Solving Equations (12) and (13) for the unknowns t_n^Q, X_n^Q , yields:

$$t_n^Q = \frac{\alpha t_n^G - w t_n^C}{\alpha - w} \quad (14)$$

$$X_n^Q = X_0 + w \left(\frac{\alpha t_n^G - w t_n^C}{\alpha - w} - t_n^C \right) \quad (15)$$

For all cycles, the point Q_n can be calculated by using equations (14) and (15).

If there is no probe vehicle for a cycle, since the shockwave line for n th cycle by using the formula derived is independent from $(n-1)$ th and $(n+1)$ th cycles, it is assumed that it has the same shockwave line as the previous one. Once the shockwave speeds are estimated, the queue dynamics can be predicted.

APPLICATION TO SIMULATION DATA

The application of the formulation and its performance are tested in simulation software VISSIM. A single one-lane link of 1 km is created to generate the necessary data. A traffic signal with 45 second green and 45 second red phases is created at the end of the link. All vehicles are passenger cars with desired speed of 50 km/h. Vehicles are assumed to be stopped when their speed drops below 5 km/h. Simulation resolution is set to five time steps per second. All the parameters are kept at the default values built within VISSIM. Simulation is run for 1800 seconds (20 cycles). Vehicle input rates of 900 vph and 1,050 vph are simulated selected in order to have both under-saturated and over-saturated conditions.

In order to check the performance of the formulation derived to determine the critical points to find the maximum queue length, several scenarios are considered. These scenarios are created by varying the probe vehicle penetration rate (10%, 20%, and 30%). Simulation was run five times for each penetration rate. Figure 4, Figure 5 and Figure 6 show estimated critical points and shockwave lines together with the location of all vehicles when they join back of queue determined from simulation. In these plots, the signal is located at a distance of 1000 m. These figures show all the vehicles when the first time they come to a stop, the estimated critical points, shockwave lines and the location of the probes used in finding these critical points. Figure 4 is for a scenario in which probe vehicle percentage is 30% whereas in Figure 6 the probe percentage is 10%, a relatively low value as compared with previous studies. As it can be observed in Figure 5 and Figure 6, there are cycles in which no probe vehicles are observed. In both scenarios, although no probe vehicles are observed for some cycles (e.g., cycle 6 in Figure 5, cycles 12 and 19 in Figure 6) the estimated critical points are reasonably close to the data obtained from simulation.

To apply the formulation explained in the previous section, each cycle needs to be tested whether it is in under-saturated or over-saturated condition. In order to do this, the following test is performed: If the space coordinate of $R_{n+1} > X_0$ (space coordinate of stop line which is 1000 in this study), cycle n is determined to be under-saturated, otherwise it is over-saturated. This test is realistic since the point R_{n+1} cannot be after the stop line.

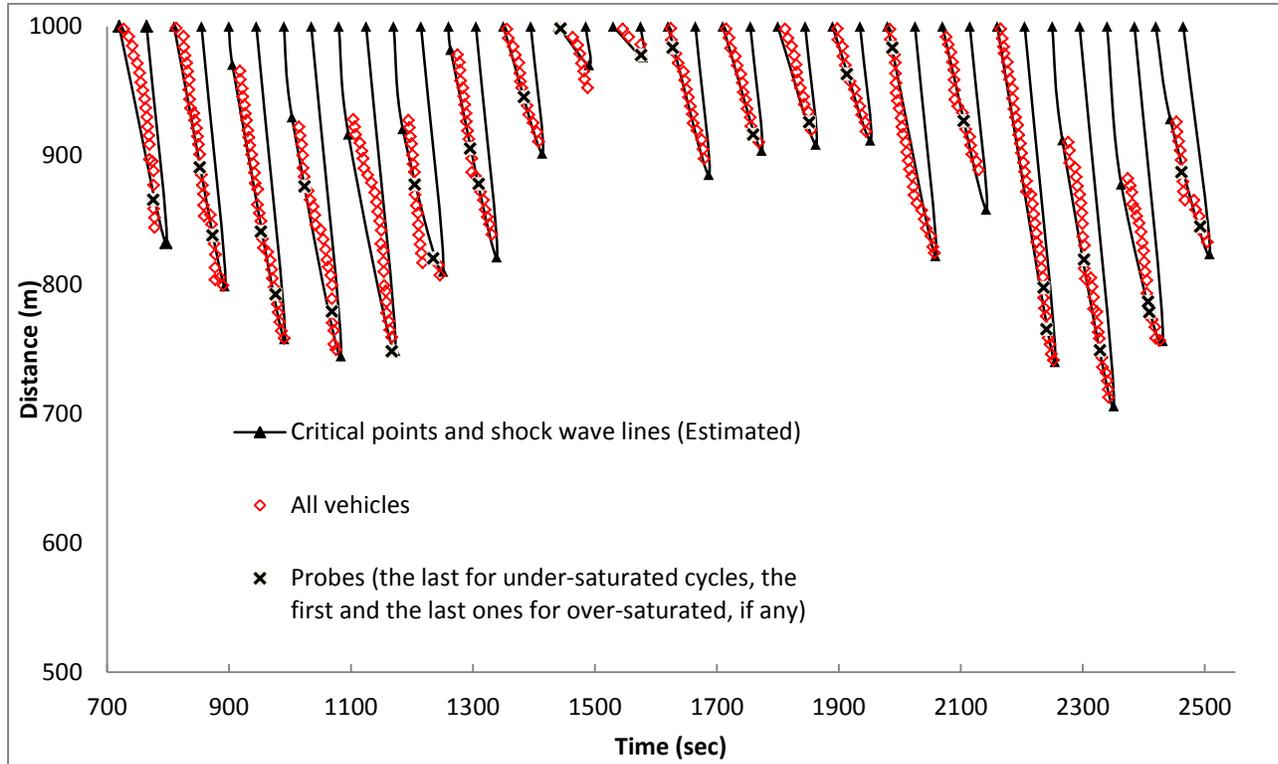


Figure 4: Estimated back of queue and shockwave lines with all vehicles (30% probe).

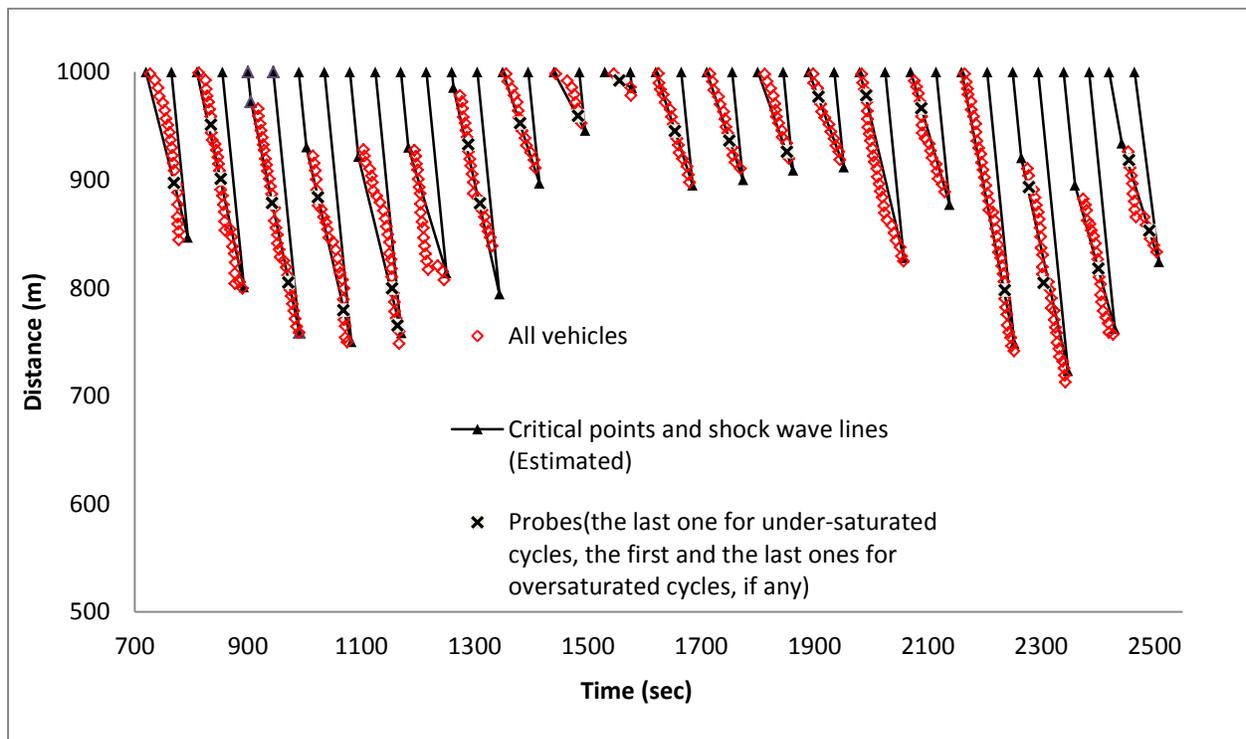


Figure 5: Estimated back of queue and shockwave lines with all vehicles (20% probe).

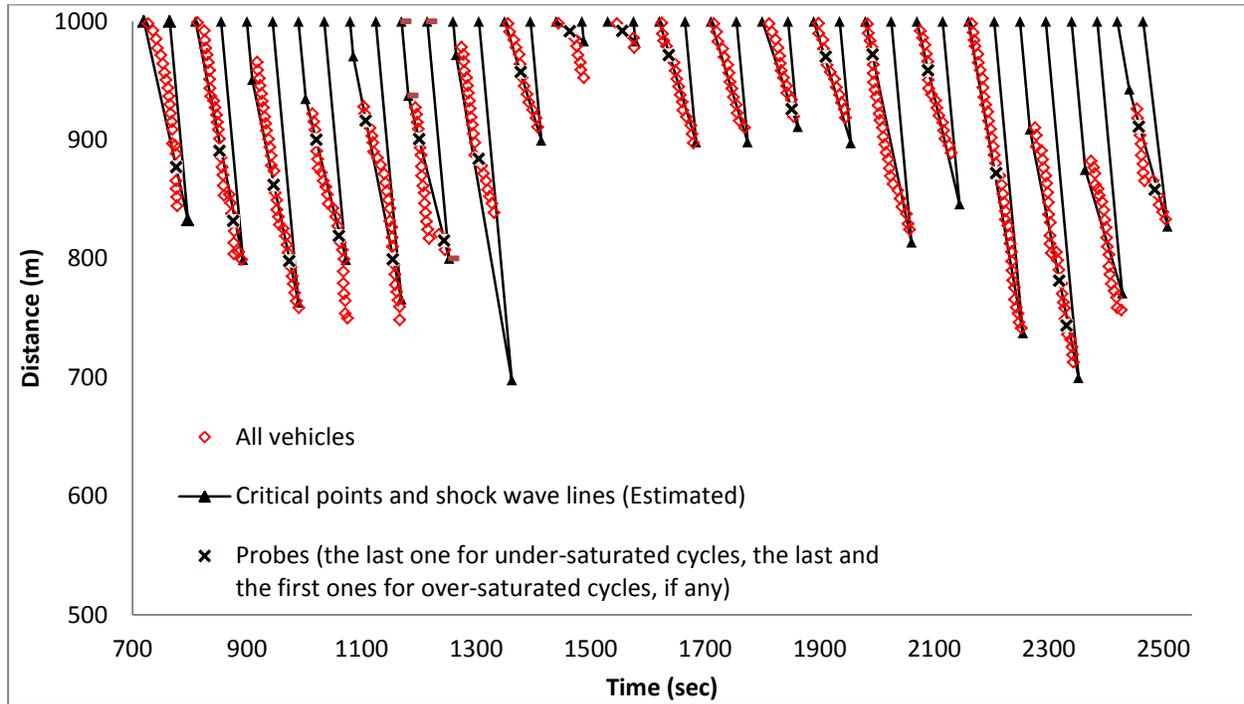


Figure 6: Estimated back of queue and shockwave lines with all vehicles (10% probe).

Summary statistics corresponding to three probe levels (i.e., 10%, 20%, and 30%) are shown in Table 1. The VISSIM simulation model is run only once to get the vehicle trajectories. For each of the three levels, the probe vehicles are selected randomly among all vehicles to generate five replicas. The critical points are predicted for 20 cycles in each replica. Therefore, there are 100 queue predictions for each probe level. The queue length (in meters) is calculated by using critical point Q_n from the stop bar and is compared with the maximum queue obtained from simulation for each cycle in each replica. The error is calculated by subtracting the estimated value from the actual value and dividing the result by the actual value. Table 1 shows the average and standard deviation of the error along with the percentage of cycles (out of 100) those have errors larger than $\pm 10\%$ and $\pm 20\%$. It can be concluded from Table 1 that as probe percentage increases, the error decreases.

Table 1: Summary Statistics for Queue Length Estimation at Different Probe Levels

Probe Level	Avg	Stdev of % Error	Percentage of samples with errors greater or less than given thresholds			
			< -10%	> 10%	< -20%	> 20%
10%	23.6%	10.3%	27%	11%	13%	7%
20%	11.0%	11.6%	29%	12%	8%	7%
30%	9.0%	32.1%	26%	8%	4%	4%

To assess the performance of the formulation under heavier congestion and longer over-saturation conditions, the input demand is increased to 1,150 vph whereas other VISSIM parameters are kept the same. Similar to the previous graphs, Figure 7 shows the evolution of the back of the queue profile as well as the predicted profile for a scenario when the probe percentage is 10. As it can be observed, there is a more substantial growth in queue, in comparison to the previous graphs. In the scenario with 10% probes shown in Figure 7, at least one probe data point is observed in each cycle. The estimated queue profiles based on the limited probe data follow the true profile reasonably well.

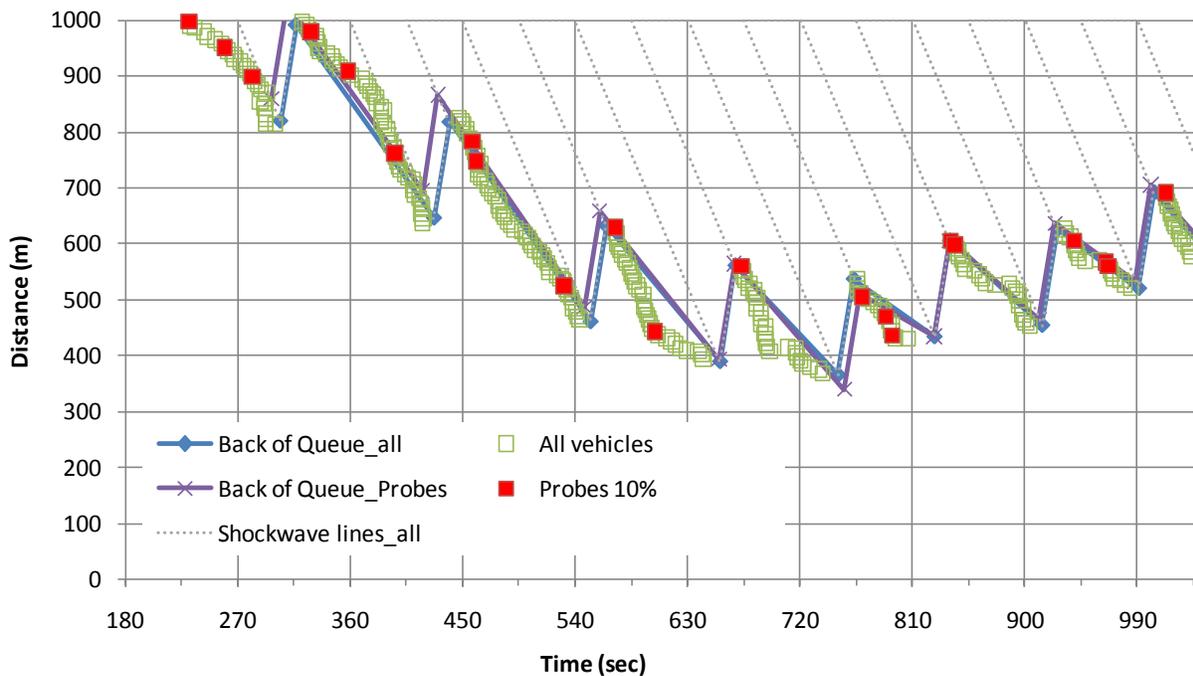


Figure 7: Back of queue estimated with all vehicles and 10% probes (input volume 1150 vph).

The results graphically shown so far in Figure 7 (and in previous graphs) are for a single random sample of probe vehicles at a given rate. To investigate the sample variability, Figure 8 shows the back of queue profiles estimated from 10 different replicas generated at 5% probe rate when the vehicle input is 1,150 vph. The critical points for the true profile are indicated by diamond markers. Overall, most of the estimated profiles are following the true profile reasonably close. In several of the replicas, there is some deviation from the true answer, especially for cycles 4 and 5 where there is a sudden jump in queue size. After examining the data, it is observed that these differences tend to get larger when there are no probe data points nearby the critical points.

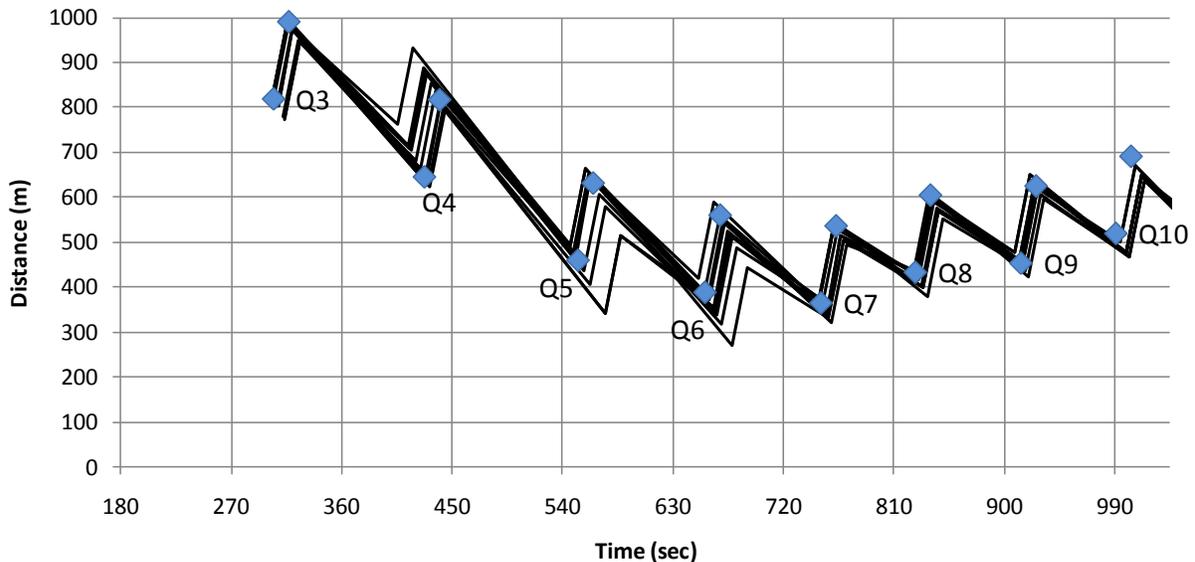


Figure 8: Ten different back of queue profiles estimated with randomly selected 5% probes.

To gain further insights into the impacts of probe vehicles on the queue prediction accuracy, summary statistics corresponding to four different probe levels are tabulated in Table 2 for the scenario where input flow rate is set to 1,150 vph. The VISSIM model is run only once to generate vehicle trajectories. For each one of the four levels (i.e., 5%, 10%, 15%, and 20%), the probe vehicles are selected randomly among all vehicles to create 20 different replicas. The back of the queue is predicted for a total of 9 cycles in each replica. Therefore, there are 180 queue predictions ($20 \times 9 = 180$) for each probe level. The queue is measured in meters from the stop bar

(i.e., the vertical coordinate of point Q_n measured from stop bar) and is compared to the predicted value in each cycle in each replica. The error is found by subtracting the estimate from the actual and by dividing the result by the actual queue length. The average and standard deviation of the error are shown in Table 2 along with the percentage of cycles (out of 180) that exhibit errors larger than $\pm 10\%$ and $\pm 20\%$. Overall, it can be observed that error decreases as probe percentage is increasing as expected. It is found that the predicted queue lengths are longer than the actual values since the average error is negative. This is perhaps an artifact of using a single VISSIM run to generate the input data. Additional runs at the same input flow level and at different flow levels can be conducted in the future to investigate the performance of the method more comprehensively.

Table 2: Summary Statistics for Back of Queue Prediction at Different Probe Levels (input volume 1150 vph)

Probe Level	Avg	Stdev of % Error	Percentage of samples with errors greater or less than given thresholds			
			< -10%	> 10%	< -20%	> 20%
5%	-4%	10%	20%	6%	6%	2%
10%	-5%	7%	18%	1%	4%	1%
15%	-2%	6%	8%	3%	0%	1%
20%	-2%	5%	4%	1%	1%	0%

DISCUSSION OF RESULTS

As illustrated in the previous section, the proposed method is effective in estimating the queue dynamics from limited probe vehicle observations for both under-saturated and over-saturated conditions. The methodology can be particularly useful, even with a relatively small number of probe observations, to evaluate the performance of signalized intersections experiencing significant congestion or over-saturation. As probe vehicle data become available, it is important to develop similar methodologies to make the best use of such data.

The methodology shown here provides a much richer understanding of congestion and system performance than just travel times and delays that are typically estimated from probe data. For example, the evolution of queues over time (e.g., peak hours) can be estimated from limited probe data. Since, signal timing data (e.g., phase and cycle lengths) are collected by some traffic operations centers, the formulation presented here can be used to make the best use of the probe data.

The methodology can be extended to real-time applications to predict the back of queue at every cycle (e.g., at the end of red phase) which can help better optimize signal timing. For example, predicting how far the queue will grow for each cycle (i.e., more precisely the critical points Q_n as described in the paper) help determine the green time needed to clear the queue so that no residual queue is left in the next cycle.

In addition, the work presented here can help with the development of data collection policies from probe vehicles as the methodology shows what type of data is more useful than others to predict queues and consequently systems performance. For example, in the context of this study, only event data when vehicles join the back of the queue for the first time are utilized. Even though vehicles may make subsequent stops before departing at the stop bar (e.g., under over-saturated conditions) the data pertaining to these other events are not needed (for the methods developed here).

CONCLUSIONS

A new methodology is presented in this study to estimate the queue by using the probe vehicle data at signalized intersections for a time period without knowing the type of cycle condition (either under-saturated or over-saturated). An analytical formula is developed to estimate the queue length for under-saturated and over-saturated conditions. Firstly, by using probe vehicle data, whether the cycle is in under-saturated or over-saturated condition is tested. Once this is determined, either the formulation for under-saturated condition or the formulation for over-saturated condition is applied. The formulation for under-saturated condition estimates the maximum queue of each cycle by using only one probe data point (i.e., the last point observed in that cycle) and it does not depend on the data from other cycles. If there is no probe vehicle in a given cycle, it is assumed that it has the same shockwave speed as in the previous cycle. The performance of the formulation is evaluated on data from VISSIM simulation runs. The results show that the method is effective in estimating the queue length for both under-saturated and over-saturated conditions. For example, under the over-saturated conditions simulated here, the error is less than $\pm 10\%$ in more 90% of the cycles when the market penetration of probe vehicles is 15%.

ACKNOWLEDGEMENTS

This study was funded by the TranLIVE University Transportation Center (UTC), a consortium of five universities, funded by the US DOT's Research and Innovative Technology Administration (RITA).

REFERENCES

1. Ban, X., P. Hao and Z. Sun. Real Time Queue Length Estimation for Signalized Intersections Using Travel Times from Mobile Sensors. *Transportation Research Part C: Emerging Technologies*, Vol. 19, No. 6, 2011, pp. 1133-1156.
2. Ban, X. G., R. Herring, P. Hao and A. M. Bayen. Delay Pattern Estimation for Signalized Intersections Using Sampled Travel Times. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2130, 2009, pp. 109-119.
3. Cetin, M. Estimating Queue Dynamics at Signalized Intersections from Probe Vehicle Data Methodology Based on Kinematic Wave Model. *Transportation Research Record*, No. 2315, 2012, pp. 164-172.
4. Comert, G. Simple Analytical Models for Estimating the Queue Lengths from Probe Vehicles at Traffic Signals. *Transportation Research Part B-Methodological*, Vol. 55, 2013, pp. 59-74.
5. Comert, G. Effect of Stop Line Detection in Queue Length Estimation at Traffic Signals from Probe Vehicles Data. *European Journal of Operational Research*, Vol. 226, No. 1, 2013, pp. 67-76.
6. Comert, G. and M. Cetin. Queue Length Estimation from Probe Vehicle Location and the Impacts of Sample Size. *European Journal of Operational Research*, Vol. 197, 2009, pp. 196-202.
7. Comert, G. and M. Cetin. Analytical Evaluation of the Error in Queue Length Estimation at Traffic Signals from Probe Vehicle Data. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 12, No. Copyright 2011, The Institution of Engineering and Technology, 2011, pp. 563-573.
8. Unal, O. and M. Cetin. Estimating Queue Dynamics and Delays at Signalized Intersections from Probe Vehicle Data. *Transportation Research Board 93rd Annual Meeting*, 2014,
9. Lawson, T. W., D. J. Lovell and C. F. Daganzo. Using Input-Output Diagram to Determine Spatial and Temporal Extents of a Queue Upstream of a Bottleneck. *Transportation Research Record*, No. 1572, 1996, pp. 140-147.
10. Erera, A. L., T. W. Lawson and C. F. Daganzo. Simple, Generalized Method for Analysis of Traffic Queue Upstream of a Bottleneck. *Transportation Research Record*, Vol. 1646, 1998, pp. 132-140.
11. Lighthill, M. J. and G. B. Whitham. On Kinematic Waves. I. Flood Movement in Long Rivers. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences*, Vol. 229, No. 1178, 1955, pp. 281-316.

-
12. Lighthill, M. J. and G. B. Whitham. On Kinematic Waves. ii. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences*, Vol. 229, No. 1178, 1955, pp. 317-345.
 13. Richards, P. I. Shock Waves on the Highway. *Operations Research*, Vol. 4, No. 1, 1956, pp. 42-51.
 14. Liu, H. X., X. Wu, W. Ma and H. Hu. Real-Time Queue Length Estimation for Congested Signalized Intersections. *Transportation Research Part C: Emerging Technologies*, Vol. 17, 2009, pp. 412-427.
 15. Van Lint, J. W. C., S. P. Hoogendoorn and M. Schreuder. Fastlane: New Multiclass First-Order Traffic Flow Model. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2088, , Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 177-187.
 16. Wong, G. C. K. and S. C. Wong. A Multi-Class Traffic Flow Model - an Extension of Lwr Model with Heterogeneous Drivers. *Transportation Research Part A: Policy and Practice*, Vol. 36, 2002, pp. 827-841.
 17. Zhang, H. M. and W. L. Jin. Kinematic Wave Traffic Flow Model for Mixed Traffic. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1802, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 197-204.
 18. Geroliminis, N. and A. Skabardonis. Prediction of Arrival Profiles and Queue Lengths Along Signalized Arterials by Using a Markov Decision Process. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1934, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 116-124.
 19. Viti, F. and H. J. Van Zuylen. Modeling Queues at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1883, Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 68-77.
 20. Cetin, M. and H. Rakha. Estimating Fuel Consumption and Carbon Footprint at Signalized Intersections Using Probe Vehicle Trajectories. *Transportation Research Board 93rd Annual Meeting*, 2014,
 21. Rakha, H. A., K. Ahn, K. Moran, B. Saerens and E. Van den Bulck. Virginia Tech Comprehensive Power-Based Fuel Consumption Model: Model Development and Testing. *Transportation Research Part D-Transport and Environment*, Vol. 16, No. 7, 2011, pp. 492-503.

APPENDIX

Publications resulting from this project:

O. Unal and M. Cetin “Estimating Queue Dynamics and Delays at Signalized Intersections from Probe Vehicle Data” Presented at the *93rd Annual meeting of the TRB*, Washington, D.C., January 12-16, 2014.

M. Cetin and H. Rakha “Estimating Fuel Consumption at Signalized Intersections from Probe Vehicle Trajectories” Presented at the *93rd Annual meeting of the TRB*, Washington, D.C., January 12-16, 2014.