Model-Based Approach for Queue and Delay Estimation at Signalized Intersections with Erroneous Automated Data

S. P. Anusha[1]; A. Sharma[2]; L. Vanajakshi, A.M.ASCE[3]; S. C. Subramanian[4]; and L. R. Rilett, M.ASCE[5]

Abstract: A model-based scheme has been developed in this paper to estimate the number of vehicles in queue and the total delay, which are typical measures for characterizing a signalized intersection. The estimation was first carried out by using the input-output and queue accumulation polygon methods. These were evaluated by using data from two instrumented intersections in the city of Lincoln, Nebraska. However, the estimates did not agree well with the actual data primarily because of the errors in counts obtained from the automated detectors. Hence, an analysis of the characteristics of the errors was carried out. A model-based estimation approach using the Kalman filter was then developed. Suitable modifications were further applied to the estimation process by incorporating calibration constants for particular site/traffic conditions in the Kalman filter estimation scheme. The results obtained were promising, indicating that the scheme could be used for performance analysis of signalized intersections with automated detectors. **DOI: 10.1061/(ASCE)TE.1943-5436.0000835.**

Author keywords: Detector error; Traffic modeling; Kalman filter; Queue estimation.

Introduction

Queue lengths and average stopped delay are the two primary performance measures used as level of service surrogates and control parameters for optimal signal operations. Efficient signal operation relies on accurate real-time estimation of these performance measures. Several methods have been proposed in the literature to obtain estimates of queue lengths and delay, which can be grouped into two categories: planning-oriented and operations-oriented approaches.

Planning-oriented approaches for intersection design are commonly performed offline during the preliminary phase of the project. Typical examples include queuing theory–based and shockwave-based methods, and in general, they are developed on the basis of historical data. There have been several queuing theory–based analytical methods (Webster 1958; Akcelik 1980, 1981; Viti and Van Zuylen 2004; Strong and Rouphail 2006; HCM 2000; Mulandi and Martin 2011; Keita and Saito 2011; Cheng et al. 2012) that estimated expected queue/delay by using (1) a historical measurement of flow rate, (2) an assumed arrival distribution, and (3) an assumed signal phase duration. The main disadvantage of these queuing theory–based methods is that they do not determine the length of the section covered by the queue. The shockwave-based methods (Lighthill and Whitham 1955; Richards 1956; Liu et al. 2009) overcome this deficiency by capturing realistic queuing behavior and determining the spatial extent of the queue. These planning oriented approaches are widely used for offline signal timing optimization, level of service estimation during the planning phase, turn bay sizing, and others. However, they are not suitable for operational analysis because they require methods that can work better by using real-time data.

Operations-oriented approaches commonly use the real-time measurements of traffic characteristics and phase display information to estimate delay and queue lengths. This real-time measurements at intersections are made by using two detectors, one placed well before the stop line, called *advance detector*, and another located near the stop bar, called *stop bar detector*. These detectors provide presence-absence information from which a variety of location-based information, such as vehicle count, headway, and phase information, can be inferred. One of the simplest approaches to estimate queue from these detector data is to use the arrival and departure counts of individual vehicles, using the input-output method (May 1990). In this method, an initial count of the number of vehicles in queue between the advance and stop bar detector is first identified, to which the number of vehicles entering the section is continuously added and the number of vehicles leaving the section is continuously subtracted to get the number of vehicles in queue inside the intersection. One issue with the input-output method is that the initial queue at the start of analysis should be known, which is often difficult to obtain. Another critical limitation is its high sensitivity to count errors, (Briedis and Samuels 2010; Vanajakshi and Rilett 2004; Rene et al. 2011) which would result in an inaccurate estimation of queue. There are many reasons for errors, including count errors caused by the presence of right-turn vehicles that are detected at the advance detector but not at the stop bar detector, vehicle lane changes, missed vehicles, and/or false-positive calls from the detectors. The following extra measurements have been used in the past to provide additional information about the number of vehicles in queue to tackle the errors caused by detector errors:

[1]Research Scholar, Dept. of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India. E-mail: anushanair1@yahoo.co.in

[2]Associate Professor, Dept. of Civil Engineering, Univ. of Iowa, Iowa City, IA 52242. E-mail: anujs@iastate.edu

[3]Associate Professor, Dept. of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India (corresponding author). E-mail: lelitha@iitm.ac.in

[4]Associate Professor, Dept. of Engineering Design, Indian Institute of Technology Madras, Chennai 600036, India. E-mail: shankarram@iitm.ac.in

[5]Professor and Director, Mid-America Transportation Center, Univ. of Nebraska—Lincoln, Lincoln, NE 68583-0851. E-mail: lrilett2@unl.edu

Note. This manuscript was submitted on May 22, 2014; approved on October 30, 2015; published online on January 27, 2016. Discussion period open until June 27, 2016; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering***, © ASCE, ISSN 0733-947X.

1. Red start time: In traffic conditions in which the queue clears in the subsequent green period, it can be assumed that the queued vehicles will clear the intersection during the green period and no residual queue would remain at the start of the subsequent cycle. On the basis of this assumption, the queue length can be made zero at the start of every red, partly addressing the problem of inaccurate queue estimates. The assumption of zero queue at the start of red would be violated in cases in which the queue is not discharged in the allotted green time of a given cycle, resulting in residual queues.
2. Green start time: The green start time, average start-up lost time and the average saturation flow rate can be used to generate the departure flow profile (HCM 2000). The accuracy of the modeled flow profile depends on the accuracy of these input values because the generated flow profile is dependent on these input values.
3. Queue clearance times: The queue clearance time is the time it takes to completely clear the queue and can be estimated by using information from the stop bar detector. This calculation is based on the fact that the headway of vehicles being cleared from a queue is the minimum (queue clearance minimum headway) after the start-up lost time, and the headway between vehicles after the queue clearance phase will be higher than that. The time at which a headway greater than the queue clearance minimum headway is first observed at the stop bar detector after the start of green is the queue clearance time. By definition, queue clearance times can only be obtained for signal cycles, in which the queue clears within the subsequent green period. The estimated queue clearance time can be incorrect if there are missed vehicles and/or false-positive calls at the stop bar detector, driver delay during the queue clearance process (e.g., inattentive drivers), or an incorrect minimum queue clearance headway is used.
4. Time occupancy at the advance detector: The occupancy time at the advance detector can be used to estimate the number of vehicles in queue (Vigos et al. 2008). One concern in using this approach is that the relationship between time occupancy and the number of vehicles in the queue is site specific. Time occupancy can be used for queue estimation only for the cases in which the queue ends within the advance detector because vehicle arrivals and hence time occupancy cannot be obtained once the advance detector is fully occupied. Another concern is that the accuracy of time occupancy data is dependent on the detector polling rate. If the detector polling rate is too slow, the vehicles may not get identified, which can have a large negative effect on the accuracy of the estimated number of vehicles in the queue.

Some of the recently reported studies using the aforementioned information to estimate queue lengths and/or delays are discussed subsequently.

Delay estimation using real-time algorithms was carried out by Balke et al. (2005), Skabardonis and Geroliminis (2005), Smaglik et al. (2007), and Liu and Ma (2008). Different methodologies for queue estimation using loop detector data and signal timing data were used by Sharma et al. (2007) and Liu et al. (2009). Various other methodologies that used time occupancy data from loop detectors to determine queue lengths have also been reported (Chang et al. 2012; Qian et al. 2012). However, none of the aforementioned studies have addressed the modeling uncertainties and detector errors. There have been studies that addressed this concern by using the Kalman filter estimation scheme (Lee et al. 2013; Vigos et al. 2008; Wu et al. 2008). All of these studies used time occupancy as the measurement variable. Based on this, it is assumed that the corresponding queues are within the advance detector because time occupancy can be obtained from the advance detector under this condition only. Table 1 shows the details of these studies.

Table 1. Comparison of Kalman Filter–Based Estimation Schemes for Queue

Reference	Component	Details
Vigos et al. (2008)	Time resolution	20 s
	Estimate	Number of vehicles in the link; between two intersections
	Input	Entering vehicle flow—exiting vehicle flow
	Measurement	Time occupancy from the detector located midway between entry and exit
	Process disturbance	Assumed Gaussian
	Measurement noise	Assumed Gaussian
	Conditions modeled	Simulated data; queue within advance detector
	Remarks	The impact of various factors such as the estimation scheme parameters, traffic parameters, and geometric features were investigated. The simulated data being error free, the detector errors were not considered in the analysis
Lee et al. (2013)	Time resolution	0.5 s
	Estimate	Number of vehicles in queue on the ramp
	Input	Entering vehicle flow—exiting vehicle flow
	Measurement	Time occupancy from the mid-link detector and entry detector
	Process disturbance	Assumed Gaussian
	Measurement noise	Assumed Gaussian
	Conditions modeled	Simulated data, queue within advance detector
	Remarks	Detector errors were introduced in the simulation. However, the influence of the detector errors on the estimation scheme was not explicitly analyzed
Wu et al. (2008)	Time resolution	20 s
	Estimate	Number of vehicles in queue on the ramp
	Input	Entering vehicle flow—exiting vehicle flow
	Measurement	Time occupancy from loop detector at entry
	Process disturbance	Calibrated constants based on observed queue data
	Measurement noise	Assumed Gaussian
	Conditions modeled	Field data; queue within advance detector
	Remarks	Adjustment factors were provided to balance the miscounting effects of the loop detectors. Site-specific constants to account for the process disturbance were included. The errors in the measurement variable namely time occupancy were not analyzed

It can be observed that in most of the previous studies, time occupancy at the advance detector and/or the midlink detector was used to estimate queue. However, in this study, it was found that time occupancy was not the best indicator of the number of vehicles in the queue as will be discussed in the "Methodology" section. Another key observation from the literature review was that there are only limited reported studies that had explicitly incorporated the detector errors while estimating queue and delay. Those limited studies that have incorporated the detector errors (Wu et al. 2008) have developed calibration constants only for the process disturbance. However, the measurement variables are also prone to measurement noises caused by the detector errors, which were not taken into account.

The present study developed calibration constants for the process disturbance and the measurement noise for different study sites/traffic conditions and used them in the estimation scheme. Most of the reported studies used simulated data for corroborating the results. In the present study, the developed estimation schemes were evaluated by using field data. The vehicle arrival rate in the study sites were low (an average of 200–320 vehicles per hour per lane); hence, it was assumed that the queues formed will be within the advance detector, and the queue formed for every cycle will clear in the subsequent green period (under saturated condition). In this study, queue is defined as the number of vehicles affected by red in an intersection approach at any given time. Delay was estimated in terms of aggregate delay and is defined as the total delay encountered by the vehicles in queue during the considered time interval.

Two estimation schemes have been proposed in this study, which use the erroneous loop detector data for queue estimation. The common and unique features of these estimation schemes are listed in Tables 2 and 3, respectively. The main difference between the two estimation schemes is that the calibration constants were assumed to be Gaussian in Estimation Scheme 1, whereas actual calibration constants based on the field data were used for the process disturbance and the measurement noise in Estimation Scheme 2.

Data Collection

The proposed estimation schemes were evaluated by using data collected from Lincoln, Nebraska, at two test sites: 17th Street and G Street (17&G) and 27th Street and U.S. Highway 6/Cornhusker Highway (27&Cornhusker Highway). The northbound (NB) approach of 17th Street, which has an average daily traffic (ADT) of 10,710 vehicles, and the eastbound (EB) approach of 27th Street, which has an ADT of 33,485 vehicles, were considered. Both the NB 17th Street approach and the EB 27th Street approach had three lanes and were equipped with monitoring equipment that could collect detector actuation data at the advance detector and the stop bar detector on a lane-by-lane basis. The signal phase information was also collected for the through approaches. The advance detectors are located in each lane at 90 and 100 m behind the stop line for the 17&G and 27&Cornhusker intersections, respectively. The stop bar detectors are located approximately 5 m upstream of the stop bar detector. The status of detector actuations and signal were collected by using a digital input-output (I/O) device that was polled by the central server at 1-s intervals. The data were stored in a structured query language (SQL) server. Fig. 1(a) shows the hardware used on the NB approach of 17th Street at the 17&G test site. A human machine interface (HMI) tool was developed, which was used to overlay the detector status information next to the video stream from the site. A screenshot of the same is shown in Fig. 1(b). This tool was used for visual verification of the data and to assess the performance of the detectors at both the test approaches. Data were collected in the right- most lane at peak and off-peak traffic conditions for both approaches. These data were used to evaluate the performance of the two estimation schemes developed in this study.

Table 2. Common Components of Estimation Schemes 1 and 2

Component	Details
Time resolution	1 s
Estimate	Number of vehicles in queue; total delay
Input	Vehicles joining the queue as a function of time Vehicles discharged from the queue as a function of time
Measurement	Number of vehicles in queue was estimated by using the queue clearance time from stop bar detector and signal timing information
Conditions modeled	Field data; queue clearing within green period and queue within advance detector traffic conditions

Table 3. Unique Components of Estimation Schemes 1 and 2

Model parameters	Estimation Scheme 1	Estimation Scheme 2
Process disturbance	Assumed Gaussian	Calibrated constants based on observed relative flow data
Measurement noise	Assumed Gaussian	Calibrated constants based on observed queue data

Basic Approaches for Queue Estimation

Queue estimation using basic approaches, namely, input-output method and the queue accumulation polygon approach (HCM 2000), were carried out. First the input-output approach was attempted and cumulative arrivals and departures were plotted. Fig. 2(a) shows a sample plot for the northbound lane C (NC) of 17&G intersection for a selected day. It can be seen that the curves are diverging over time, indicating an accumulating queue in the section, with 1,134 vehicles in queue at the end of the analysis period. However, for the selected section length of 90 m, and assuming an average vehicle length of 4 m with 1 m for gap, the maximum number of vehicles that can be accommodated in the section is 18. This clearly indicates a continuous undercounting or overcounting in the data. Fig. 2(b) shows an enlarged view of a 1-h period input-output analysis result, clearly demonstrating this problem. The solid line shows the input-output result, and the corresponding actual queue extracted manually by observation from HMI is shown by the dotted line. It can be seen that the number in queue is continuously growing instead of the growing and clearing pattern caused by red and green intervals. Video observation showed one of the reasons for this anomaly to be the count errors caused by the right-turn vehicles activating the advance detector but not activating the stop bar detector by taking the turn before the stop bar detector. Thus, it was concluded that for the automated data collected in this study, because of the associated errors, the input-output method is not a suitable option for queue estimation.

The queue accumulation polygon approach, which is another basic approach for queue estimation, was attempted next. This method uses the vehicle counts from the advance and stop bar detectors as a function of time and the observed signal timing

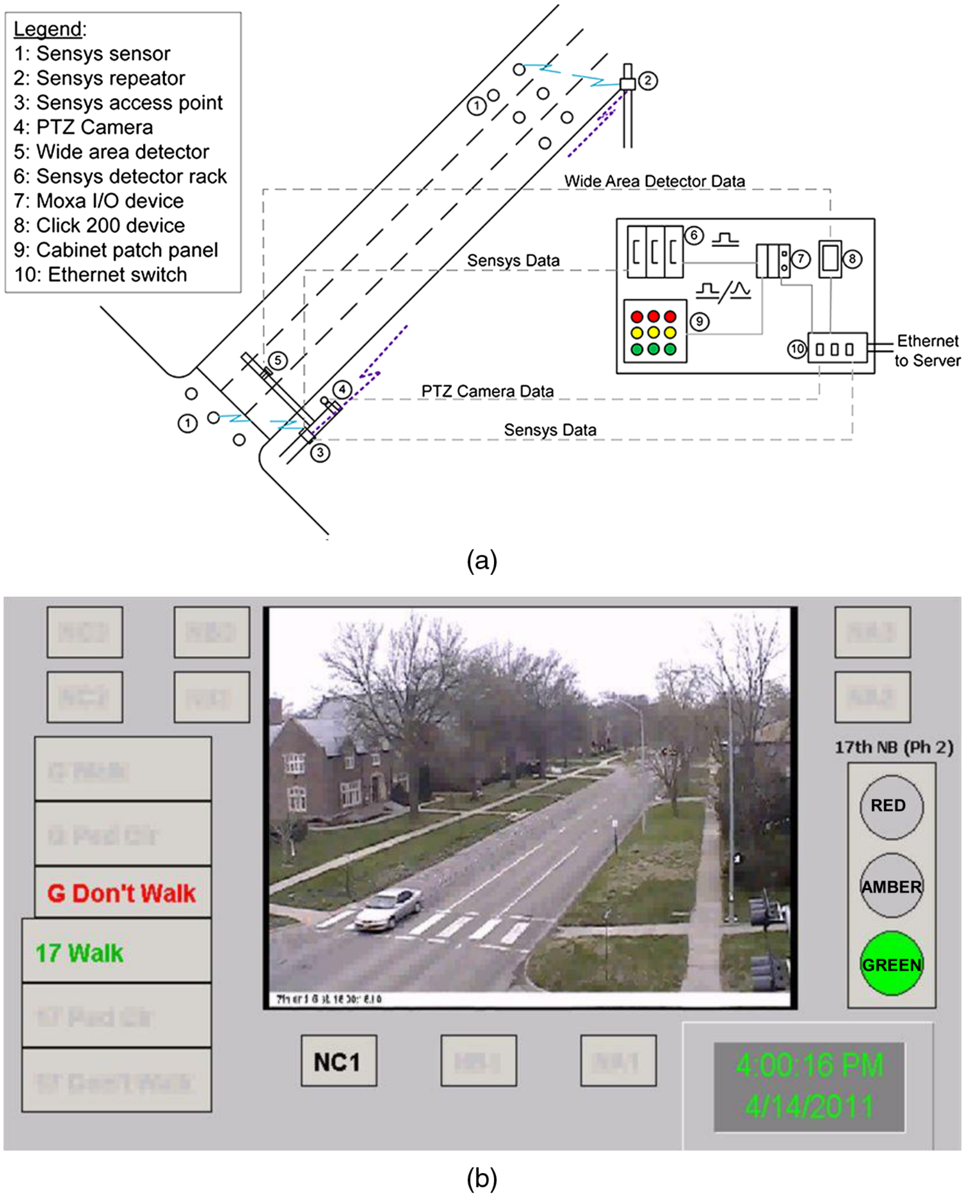

Fig. 1. Data collection schematic of the 17&G intersection

information to estimate the corresponding number in queue. The method uses the advance detector count data to develop the arrival section of the queue polygon. The maximum queue size is assumed to occur at the end of the red period. During the green interval, the number of vehicles in the queue will decrease at a rate equal to the difference between the arrival rate of vehicles at the back of queue and the departure rate of vehicles from the front of the queue. Fig. 3 shows a sample queue polygon showing the queue formation during red and the clearing of the queue during green.

Also, for the data being considered in this study, the queue formed during a red period clears in the subsequent green period. Hence, the queue at the end of any green should be zero. On this basis, the queue accumulation polygon was implemented, with a correction to ensure that the queue starts from zero at the start of every red period. In this manner, the large errors as shown by the input-output method [Fig. 2(b)] were eliminated. Fig. 4 shows the corresponding result obtained for 17&G intersection in peak traffic condition for a 1-h period starting at 8:00 a.m. The solid line indicates the estimated number of vehicles in queue, and the dotted line shows the actual queue that was obtained from manual observation. It can be seen that the estimated number in queue is not accumulating over time. However, the estimated numbers become negative in certain time intervals. To illustrate this, an enlarged view for a 5-min period is shown in Figs. 5. The queue length from 8:01:00 to 8:02:25 a.m. and from 8:03:45 to 8:04:00 a.m. have negative values. In addition, the queue failed to clear during the green interval at 8:05:25 a.m., which is not expected under the traffic condition being studied. The main causes of these errors are the wrong exit detection once the queue was cleared and the inaccuracy in the count data obtained from the advance and stop bar detectors. Hence, an analysis was carried out to identify the detector errors by comparing the loop detector data with the HMI as discussed in the following section.

Detector Error Analysis

From the previous discussion, it is clear that the accuracy of count data plays a major role in queue estimation. Hence, the accuracy of

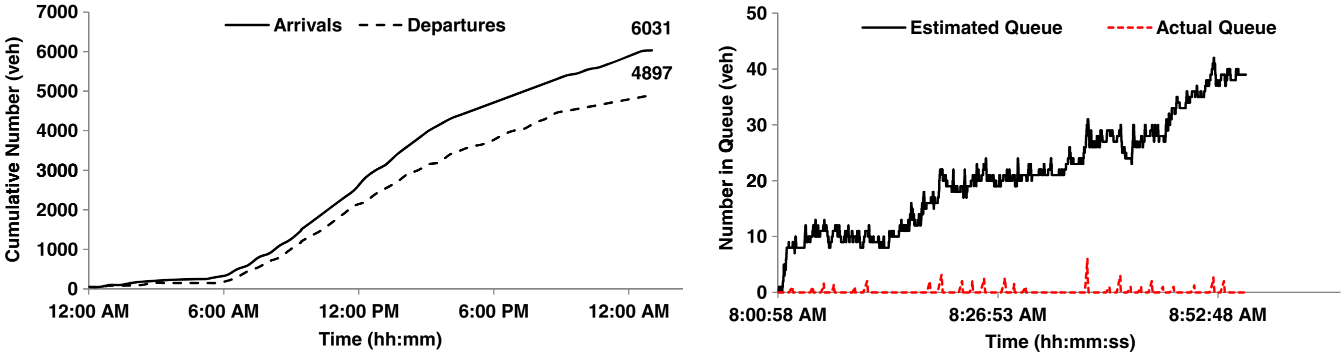

Fig. 2. Shortcomings of basic input output method for queue estimation

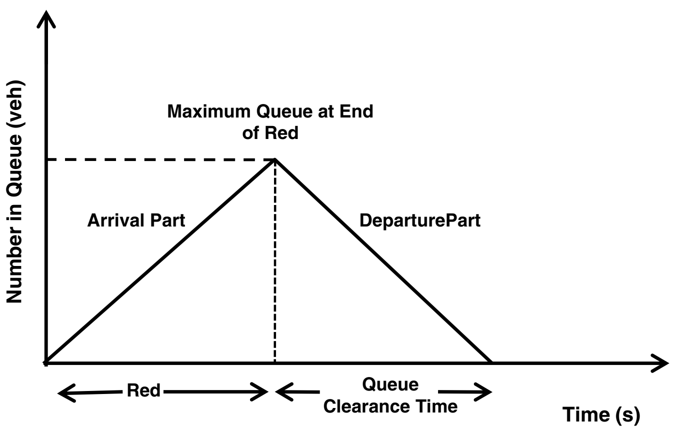

Fig. 3. Queue polygon

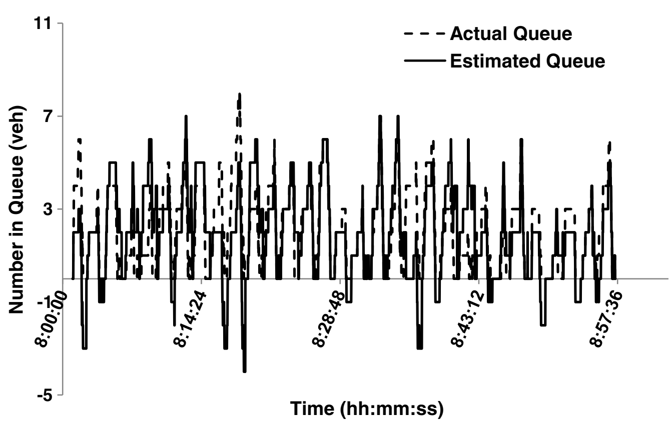

Fig. 4. Shortcomings of queue accumulation polygon approach for queue estimation

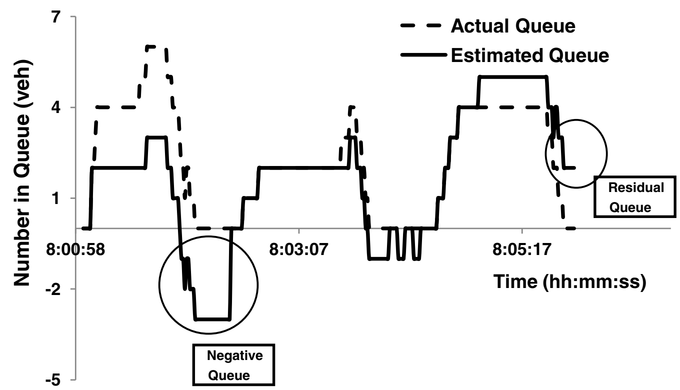

Fig. 5. Zoomed-in view of 0–5 min of Fig. 4

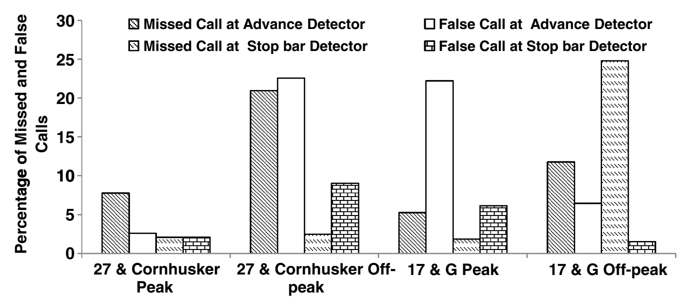

Fig. 6. Summary of missed and false-positive calls at advance and stop bar detectors

the count data obtained from the loop detector was checked by comparing them with those manually extracted by using the HMI. Data from the advance and stop bar detectors of both the intersections in peak and off-peak conditions were analyzed for missed calls and false-positive calls and are shown in Fig. 6. It can be observed that there is a significant amount of detector errors at both intersections.

These count errors will affect the accuracy of the entry and exit counts, which in turn will affect the estimated number in queue at any instant. Thus, suitable estimation schemes that can provide reliable estimate of queue and delay while using such detector data need to be identified. Model-based approaches that use recursive estimators are reported to be capable of incorporating noise in data and capturing the variations in the traffic (Dailey 1997) and hence were chosen in the present study. Two model-based estimation schemes that used erroneous loop detector data were proposed: (1) without explicit analysis of the detector errors; and (2) with explicit analysis of the detector errors. The following section describes the details of the proposed methodology.

Methodology

A model-based approach along with a Kalman filter–based estimation scheme was adopted in this study to estimate queue and delay.

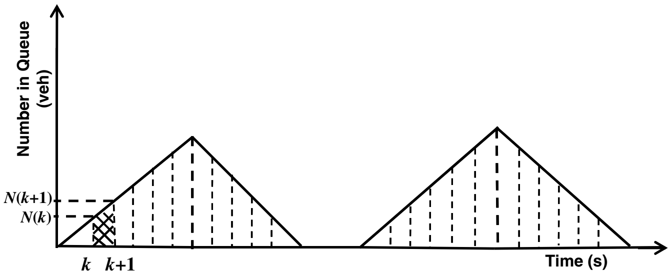

Fig. 7. Delay estimation from queue

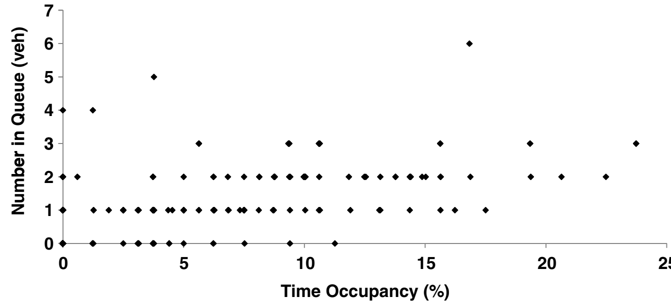

Fig. 8. Relation between time occupancy and queue at the advance detector

The basic equations to determine the number of vehicles in queue and the total delay were based on the conservation of vehicles and queue accumulation polygon approach. The conservation of vehicles states that the number of vehicles in queue at the next instant of time is equal to the sum of the number of vehicles in queue at the current instant of time and the difference between the number of vehicles that have joined and left the queue during this time interval. Eq. (1) presents this mathematically, which can be used to estimate the number of vehicles in queue at the next instant of time $N(k+1)$

$$N(k+1) = N(k) + [N_{\text{entry}}(k, k+1) - N_{\text{exit}}(k, k+1)] \qquad (1)$$

where $N(k)$ = number of vehicles in the queue during the current time instant, $N_{\text{entry}}(k, k+1)$ and $N_{\text{exit}}(k, k+1)$ = number of vehicles that have joined and left the queue, respectively, during the time interval from k to $k+1$.

The queue polygon approach uses the number of vehicles in queue estimated by the conservation equation to develop a queue polygon as shown in Fig. 7. The area of the queue polygon between two time instances gives the total delay encountered by those vehicles in queue during that time interval.

Eq. (2) presents the queue polygon delay equation from which the total delay for the time period $(k,\ k+1)$, can be obtained

$$t_d(k, k+1) = \int_{kT}^{(k+1)T} N dt = \left[\frac{N(k) + N(k+1)}{2}\right] h \qquad (2)$$

where h = duration of time between the instants k and $k+1$. By substituting for $N(k+1)$ from Eq. (1), the expression for total delay becomes

$$t_d(k, k+1) = h\, N(k) + \frac{h}{2}[N_{\text{entry}}(k, k+1) - N_{\text{exit}}(k, k+1)] \qquad (3)$$

Eqs. (1) and (3) can be used to estimate queue and delay, provided that the initial number of vehicles inside the section and the initial delay are known. However, these are not available in typical field applications. To overcome this limitation, the Kalman filter (KF), an estimation tool, was used with an assumed initial number of vehicles in queue and initial delay. The KF is a recursive algorithm that is widely used for estimation and prediction. It is usually applicable to system models/realizations that can be written in the state-space form. The linear KF is used for the analysis if the equations are linear in nature. Because the aforementioned model formulations are in state-space form and linear in nature, linear KF was used. The basic requirements of the KF estimation scheme are the state equations and the measurement equations. A system can be considered as an entity that produces outputs corresponding to inputs provided to it. The characteristics of the system are described by the state variables. State variables are those whose knowledge would completely characterize the behavior of the system. The state equation describes the evolution of the state variables with the system inputs. The measurement equation provides a relationship between the state variables and the measurement variables. Measurement variables are those that can be obtained directly or inferred from those parameters that can be measured from the field. The following section describes the methodology used for determining the measurement variable in this study.

As mentioned in the literature review, previous studies that have estimated queue by using Kalman filter estimation have used time occupancy as the measurement variable. Hence, the relation between time occupancy and number in queue was checked first. Fig. 8 shows the time occupancy with number in queue plot, and it can be observed that, with the current polling rate of detector and low volume conditions, the time occupancy and number of vehicles in queue are not exhibiting a high correlation. Hence, another measurement variable, namely, the queue clearance time, was identified in this study. The methodology adopted to obtain the queue clearance time from the signal information and the detector data is explained subsequently.

The queue clearance time, which is the time from start of green for the queue to completely clear from the stop bar location area, was determined by using the stop bar detector actuations and signal timing information. When the signal turns red, the vehicles arriving at an intersection stop, and a queue is accumulated at a rate equal to the arrival flow rate. At the onset of green, the queued vehicles will leave the intersection at saturation headway after an initial start-up lost time. The queue reduction rate is equal to the saturation flow rate minus the arrival flow rate. The queue is completely discharged at the queue clearance time and remains zero until the start of red of the next cycle. Because the arrival rate in the study stretch was low, all the vehicles arriving in red and forming a queue would be completely cleared during the green period. The queue clearance time was identified when the headway, after the initial start-up lost time, becomes greater than an assumed queue clearance headway (3 s was used in this study according to data). The queue polygon was then constructed assuming zero queue at the start of the red, with the queue uniformly increasing to a maximum queue value at the start of green and then uniformly decreasing to a value of zero at the queue clearance time, as shown in Fig. 3. In this study, queue measurements were back-calculated by using the queue clearance times and phase information. For example, the maximum number of vehicles in the queue was obtained by multiplying the duration of queue clearance time with the difference between the saturation flow rate and the arrival flow rate. After obtaining all the critical points (e.g., queue clearance time and maximum

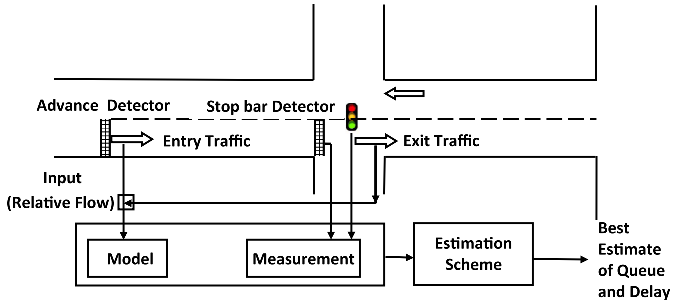

Fig. 9. Application of KF estimation to the traffic system

number of vehicles in queue at end of red), the polygon can be completed, and by using this, the number of vehicles in queue at any desired instant of time can be determined.

Estimation Scheme Using Kalman Filter

For the present study, the state variables were the number of vehicles in queue and total delay. The input to the traffic system was the relative flow (the difference between the input flow and the output flow), and the measurement variable considered was the number of vehicles in queue that was obtained by using the known queue clearance time and signal timing information. Fig. 9 illustrates the complete application of KF to the traffic system under study. The input from the traffic system, which is the relative flow, will provide the estimate of queue and delay on the basis of the mathematical model. This estimate will be corrected by the KF by using the measurement to give the best estimate of queue and delay.

In general, the state equation and the measurement equation are represented in the state-space form as

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{b}\,u(k) + \mathbf{w}(k) \tag{4}$$

$$y(k) = \mathbf{c}\cdot\mathbf{x}(k) + v(k) \tag{5}$$

where $\mathbf{x}(k+1)$ = vector of state variables at the $(k+1)$th instant of time; u and y = measured input and output, respectively; and $\mathbf{w}$ and v = process disturbance and measurement noise, respectively, that are assumed to be Gaussian with zero mean.

In the present study, these variables are as follows:

$$\mathbf{x}(k+1) = \begin{bmatrix} N(k+1) \\ t_d(k+1) \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 1 & 0 \\ h & 0 \end{bmatrix},$$
$$\mathbf{b} = \begin{bmatrix} 1 \\ \frac{h}{2} \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The aforementioned realization of the system was then checked for its controllability and observability by forming controllability and observability matrices and finding their rank. The controllability and observability matrices for this realization were obtained as

$$\mathbf{C} = \begin{bmatrix} 1 & 1 \\ \frac{h}{2} & h \end{bmatrix} \qquad \mathbf{O} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

The rank of the controllability matrix and observability matrix are 2 and 1, respectively, indicating that this realization of the system is completely controllable but not completely observable. But, even though the realization of the system was not completely observable, the evolution of the unobservable state is stable. Hence, the realization is detectable, and KF can be used with this realization.

A recursion of the KF estimation algorithm begins with estimating the a priori value of the state and its corresponding error covariance and ends with the estimation of the a posteriori value of the state and its error covariance. The best estimate of the state variable is $\hat{\mathbf{x}}^{+}(k+1)$ with an error covariance of $\mathbf{P}^{+}(k+1)$. In between, the KF calculates a quantity called *Kalman gain* $[\mathbf{k}(k+1)]$ for weighing the difference between the actual measurement $y(k+1)$ and the predicted measurement, $\mathbf{c}\cdot\hat{\mathbf{x}}^{-}(k+1)$. For the present study, the state variables were the number of vehicles in queue and total delay. The input to the traffic system was the relative flow (the difference between the entry flow and the exit flow), and the measurement variable considered was the number of vehicles in queue that was obtained by using the known queue clearance time and signal timing information. The flowchart in Fig. 10 illustrates the execution of the recursive KF scheme for the traffic system. $\mathbf{Q}$ = process disturbance covariance; R = measurement noise covariance; $\hat{\mathbf{x}}^{-}(k)$ = a priori estimate of the state at the kth instant of time; and $\hat{\mathbf{x}}^{+}(k)$ = a posteriori estimate of the state at kth instant of time.

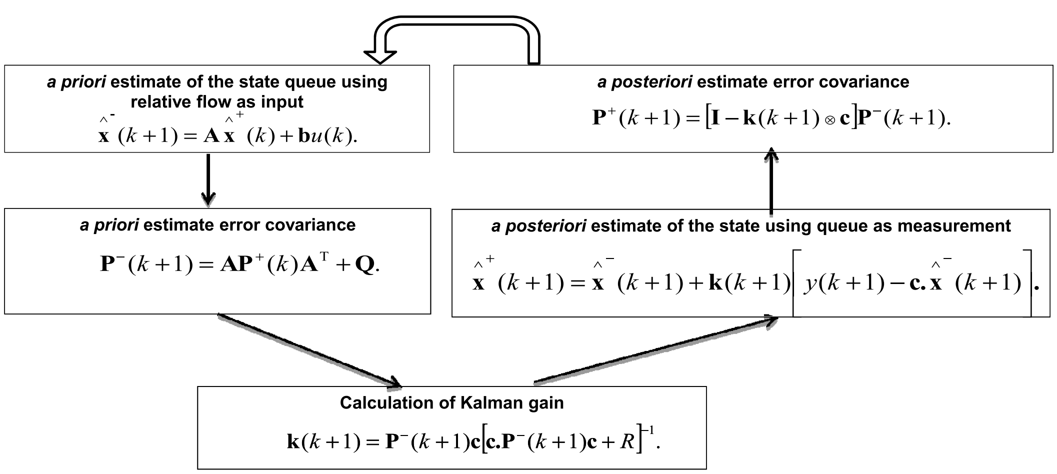

Fig. 10. Flowchart representation of KF estimation algorithm

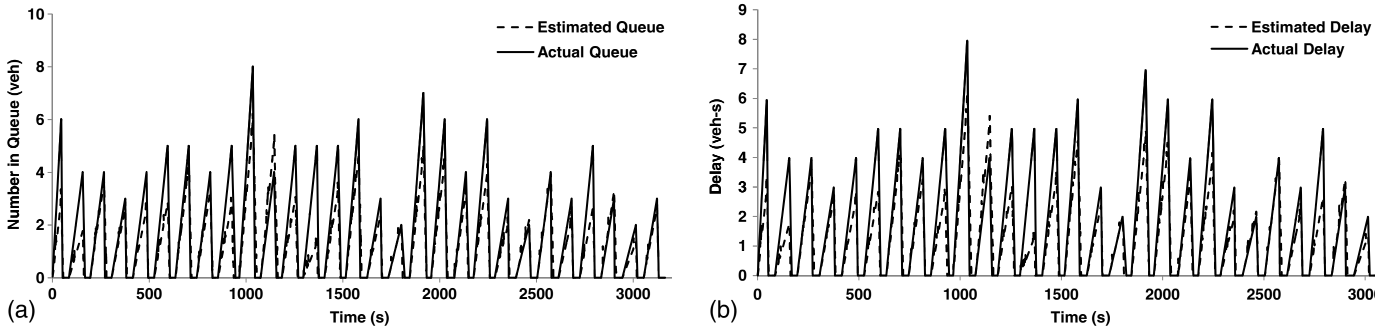

Fig. 11. Performance of KF estimation without explicit analysis of detector errors

Corroboration of the KF Estimation Scheme

The aforementioned estimation scheme was implemented and analyzed for all locations. A sample result of queue and delay estimation in 27&Cornhusker intersection during peak traffic (6:00 to 7:00 p.m.) is shown in Figs. 11(a and b). These figures show that the estimated queue matches with the actual queue, indicating the accuracy of the KF estimation scheme.

The errors in the estimation were quantified by using the RMS error (RMSE) given by

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[N_{\mathrm{act}}(k) - N_{\mathrm{est}}(k)]^2} \qquad (6)$$

where $N_{\mathrm{est}}(k)$ and $N_{\mathrm{act}}(k)$ = estimated and actual number of vehicles, respectively, in queue or delay during the kth instant, with N as the total number of time instances. The RMSE is a measure of effectiveness, which is often helpful when different methods applied to the same set of data are compared. However, there is no absolute criterion for a "good" value for any of the scale-dependent measures, as they are on the same scale as the data (Coleman and Swanson 2007). The smaller the value of the RMSE, the better the forecast obtained. The RMSE expresses the expected value of the error and has the same unit as the data, which makes the size of a typical error visible.

The RMSE obtained for queue and delay in the aforementioned case were 0.84 vehicles and 0.83 vehicle-seconds, respectively. Similar results were obtained from other sites and are shown in Table 4. Even though the estimation scheme did not explicitly take into account the properties of the detector errors, the RMSEs are lower in magnitude, indicating the accuracy of the estimation scheme. For further analysis, the KF estimation scheme was modified, in which the properties of detector errors were explicitly taken into account and is discussed in the following section.

KF Estimation with Explicit Analysis of Detector Errors

For further improvement of the proposed KF scheme, implementation was carried out by explicitly incorporating the process disturbance and measurement noise in the estimation process. The difference between the relative flow obtained from the loop detector and HMI observation was used to estimate the process disturbance. The difference between the number of vehicles in queue estimated from the queue polygon by using the loop detector data and by using HMI data was used to determine the measurement noise. The mean and variance of this process disturbance and measurement noise were calculated and were used as calibration constants of the KF estimation scheme. Table 5 lists the calibration constants for different site/traffic conditions for the data under consideration.

The state equation was suitably modified to take into account the disturbances caused by the relative flow error. To account for this, the term $m(k)$ was introduced in Eq. (4), where $m(k)$ = relative flow error (modeled as a process disturbance). Thus, $\mathbf{w}(k)$ in Eq. (4) has the form

$$w(k) = \begin{bmatrix} 1 \\ \frac{h}{2} \end{bmatrix} m(k)$$

Similarly, the mean of the measurement noise was used for $v(k)$ in Eq. (5) in the modified formulation.

The estimation scheme was implemented again using this modified formulation. Figs. 12(a and b) show a sample plot of comparison of the estimated queue and delay with the actual queue and delay, respectively, for 27&Cornhusker intersection for peak traffic condition. The corresponding RMSE in these cases were 0.73 vehicles and 0.72 vehicle-seconds, respectively. Figs. 13(a and b) show the comparison of the queue and delay that were obtained from the KF estimation scheme with and without explicit analysis of detector errors in terms of RMSE for both the intersections in peak and off-peak traffic conditions. In all the cases, the KF estimation with explicit analysis of detector errors resulted in a lower RMSE compared with KF without incorporating the detector errors. The mean of the process disturbance and measurement noise were not significantly high; hence, the changes that were caused by the explicit analysis of detector errors do not show a significant

Table 4. RMSE for Queue and Delay in Different Site/Traffic Condition

RMSE	27&Cornhusker peak	27&Cornhusker off-peak	17&G peak	17&G off-peak
Queue (vehicle)	0.837	0.628	0.251	0.207
Delay (vehicle-second)	0.834	0.627	0.250	0.229

Table 5. Calibration Constants for Different Site/Traffic Condition

Site/traffic condition	Mean of process disturbance	Variance of process disturbance	Mean of measurement noise	Variance of measurement noise
27&Cornhusker peak	−0.008	0.0793	−0.391	0.553
27&Cornhusker off-peak	0.001	0.0529	0.199	0.355
17&G peak	−0.012	0.036	−0.028	0.060
17&G off-peak	−0.016	0.322	−0.006	0.041

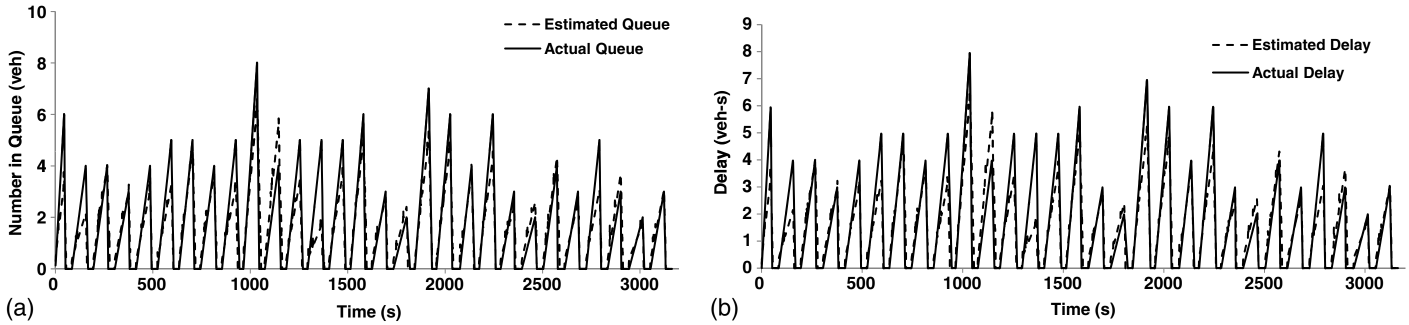

Fig. 12. Performance of KF estimation with explicit analysis of detector errors

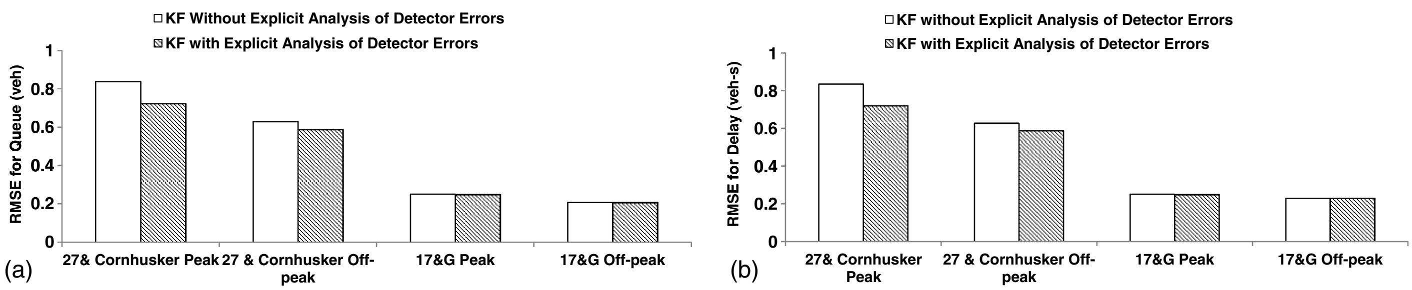

Fig. 13. Comparison of performance measures with and without explicit analysis of detector errors

difference. However, the KF estimation scheme was sensitive to the incorporation of the calibration constants of the particular site/traffic condition. Thus, it can be seen that even with erroneous data as input, the KF estimation scheme with explicit analysis of detector errors using suitable calibration constants has the potential to improve the estimation accuracy. In the aforementioned case, the calibration constants were separately calculated for each site and for peak and off-peak conditions separately, making them site specific and needs to be calculated for any new site and traffic conditions. A limited analysis was carried out to check the effect of varying traffic conditions (peak and off-peak) and different locations on the calibration constant. For this, a comparison of performance using separate calibration constants (1) for peak and off-peak conditions and for different sites, (2) for each site averaged across traffic conditions, and (3) for each traffic condition averaged across the sites was carried out. Fig. 14 shows the variation of RMSE obtained for queue for the aforementioned three cases. This figure shows that the performance of the estimation scheme by using an average calibration constant across traffic conditions for the same site was comparable to the scheme when the individual calibration constants of a particular site/traffic condition was applied. However, when the calibration constant was averaged across sites for the same traffic condition, there was an increase in the RMSE. This indicates that an average calibration constant for varying traffic conditions can be considered without much reduction in the performance of the estimation scheme. However, separate calibration constants seems to be essential for different sites.

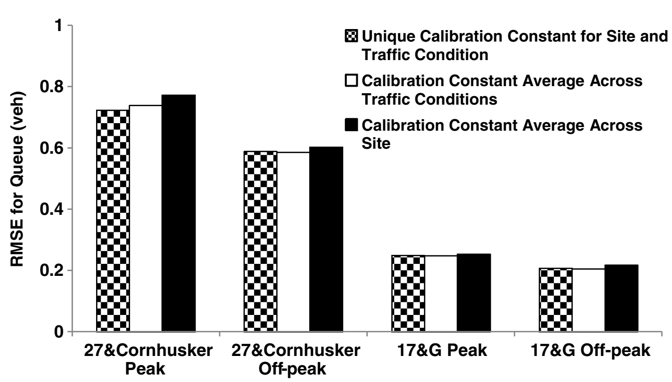

Fig. 14. RMSE for queue by applying an average calibration constant for varying site/traffic condition

In terms of frequency of recalibration, it is assumed that the performance of a detector would not deteriorate very quickly and may need recalibration only after several months. Longer duration data from multiple locations over varying traffic conditions need to be analyzed for substantiating this assumption. The identification of the duration at which recalibration will be required needs more analysis. In the current set up, whenever the performance of the estimation scheme starts deteriorating, a manual analysis can be carried out by comparing the loop detector data with the HMI observations to determine the detector errors, and recalibration can be carried out. Automatic calibration will be possible if the statistical parameters of the process disturbance and measurement noise can be obtained in an automated fashion. This would be possible if the HMI is automated, even in an offline mode.

Concluding Remarks

The data obtained from the automated sensors are used for the devising the estimation schemes for traffic operational analysis. The accuracy of the estimation scheme depends on the accuracy of the data obtained from these detectors. However, the errors caused by the detectors are inevitable; hence, suitable estimation schemes that can estimate the variables of interest while using the erroneous detector data needs to be devised. This paper presented such a reliable estimation scheme that used automated data from stop bar and advance detectors to determine the queue and delay at intersections. Data collected from automated detectors were used to determine the number of vehicles in queue at any instance by using KF estimation schemes to obtain reliable estimates under noisy detector conditions. The KF scheme was further improved by incorporating the statistical properties of the process disturbance and measurement noise in the estimation scheme. The developed schemes were tested for data collected from two intersections in peak and off-peak traffic conditions. The developed estimation scheme using the KF with explicit analysis of detector errors with suitable calibration constants was found to perform better, especially in situations with higher detector errors.

The methodology presented in this paper is for scenarios in which the road section under study is equipped with both advance and stop bar detectors. However, in the absence of a stop bar/advance detector, the corresponding flow can be estimated, and those values can be used for the queue and delay estimation. For example, if the advance detector is not available, one could use the existing model equations, along with additional equations, that relate the variables of interest to available measurements, such as occupancy, to estimate the flow through the advance detector.

As a next phase of the ongoing research, the scenario in which the queue extends beyond the advance detector, resulting in residual queues at the end of green period, is being analyzed. The formulation for this scenario was carried out by using the number of vehicles in queue as the state variable, the vehicle arrival rate as the input, and the cumulative time occupancy as the measurement variable. The initial results of the analysis are promising, and the estimation scheme is under further investigation.

The KF model can also be made self-calibrating if the statistical parameters of the estimation scheme, namely, the mean and the variance of process disturbance and measurement noise, can be obtained in an offline mode in a periodic basis (e.g., once a day or once a week). This will be possible if the HMI can be automated and the automated HMI output can then be compared with the loop detector data to determine the statistical parameters of the estimation scheme for a wide range of operating conditions, such as different weather or traffic conditions, for queue and delay estimation.

Acknowledgments

The authors acknowledge the support provided by Indo-US Science and Technology Forum (IUSSTF), Government of India, under the grant IUSSTF/JC-Intelligent Transportation Systems//95-2010/2011-12. The authors also thank Evangeline Joyce and Arul Stephen of the University of Nebraska—Lincoln for their assistance in data extraction.

References

Akcelik, R. (1980). "Time dependent expressions for delay, stop rate and queue length at traffic signals." *Australian Road Research Board, Internal Rep.*, AIR, Melbourne, VIC, Austrailia, 367–1.

Akcelik, R. (1981). "Traffic signals: Capacity and timing analysis." *Research Rep. No. 123*, Australian Road Research Board, Melbourne, VIC, Austrailia.

Balke, K., Charara, H., and Parker, R. (2005). "Development of a traffic signal performance measurement system (TSPMS)." *Rep. No. 0-4422-2*, Texas Transportation Institute, College Station, TX.

Briedis, P., and Samuels, S. (2010). "The accuracy of inductive loop detectors." *Proc., ARRB Conf.*, ARRB Group, VIC, Australia.

Chang, J., Talas, M., and Muthuswamy, S. (2012). "A simple methodology to estimate queue lengths at signalized intersections using detector data." *Proc., 91st Transportation Research Board Annual Meeting*, National Academies, Washington, DC.

Cheng, Y., Qin, X., Jin, J., and Ran, B. (2012). "An exploratory shockwave approach for signalized intersection performance measurements using probe trajectories." *J. Intell. Transp. Syst.*, 16(1), 12–23.

Coleman, C., and Swanson, D. (2007). "On MAPE-R as a measure of cross-sectional estimation and forecast accuracy." *J. Econ. Social Meas.*, 32(4), 219–233.

Dailey, D. J. (1997). "Travel time estimates using a series of single loop volume and occupancy measurements." *Proc., 76th Transportation Research Board, Annual Meeting*, National Academies, Washington, DC.

Keita, Y. M., and Saito, M. (2011). "Remote data collection for the IQA delay estimation method." *Proc., 1st Transportation and Development Institute Congress*, ASCE, Reston, VA.

Lee, J. B., Jiang, R., and Chung, E. (2013). "A Kalman filter based queue estimation algorithm using time occupancies for motorway on-ramps." *Proc., 92nd Transportation Research Board Annual Meeting*, Washington, DC.

LightHill, M. J., and Whitham, G. B. (1955). "On kinematic waves. II. A theory of traffic flow on long crowded roads." *Proc. R. Soc. London, Ser. A*, 229(1178), 317–345.

Liu, H., and Ma, W. (2008). "A real-time performance measurement system for arterial traffic signals." *Proc., 87th Transportation Research Board Annual Meeting*, Washington, DC.

Liu, H. X., Wu, X., Ma, W., and Hu, H. (2009). "Real-time queue length estimation for congested signalized intersections." *Transp. Res. Part C*, 17(4), 412–427.

May, A. D. (1990). *Traffic flow fundamentals*, Prentice Hall, Upper Saddle River, NJ.

Mulandi, J., and Martin, P. T. (2011). "Modifications to the incremental queue accumulation method for complex left turn phasing." *J. Basic Appl. Sci. Res.*, 1(3), 252–259.

Qian, G., Lee, J. B., and Chung, E. (2012). "Algorithm for queue estimation with loop detector of time occupancy in off-ramps on signalized motorways." *Transp. Res. Rec.*, 2278, 50–56.

Rene, O. S., Horowitz, R., and Varaiya, P. (2011). "Analysis of queue estimation methods using wireless magnetic detectors." *Transp. Res. Rec.*, 2229, 34–45.

Richards, P. I. (1956). "Shock waves on highway." *Oper. Res.*, 4(1), 42–51.

Sharma, A., Bullock, D., and Bonneson, J. (2007). "Input-output and hybrid techniques for real-time prediction of delay and maximum queue length at signalized intersection." *Transp. Res. Rec.*, 2035, 69–80.

Skabardonis, A., and Geroliminis, N. (2005). "Real-time estimation of travel times on signalized arterials." *Proc., 16th Int. Symp. on Transportation and Traffic Theory*, Univ. of Maryland, College Park, MD, 387–406.

Smaglik, E. J., Sharma, A., and Bullock, D. M. (2007). "Event based data collection for generating actuated controller performance measures." *Proc., 86th Transportation Research Board Annual Meeting*, Washington, DC.

Strong, D. W., and Rouphail, N. M. (2006). "New calculation method for existing and extended HCM delay estimation procedures." *Proc., 85th Transportation Research Board Annual Meeting*, Washington, DC.

Transportation Research Board. (2000). *Highway capacity manual*, National Research Council, Washington, DC.

Vanajakshi, L., and Rilett, L. R. (2004). "Loop detector data diagnostics based on conservation-of-vehicles principle." *Transp. Res. Rec.*, 1870, 162–169.

Vigos, G., Papageorgiou, M., and Wang, Y. (2008). "Real-time estimation of vehicle-count within signalized links." *Transp. Res. Part C*, 16(1), 18–35.

Viti, F., and Van Zuylen, H. J. (2004). "Modeling queues at signalized intersections." *Transp. Res. Rec.*, 1883, 68–77.

Webster, F. V. (1958). *Traffic signal settings*, Road Research Laboratory, Her Majesty's Stationery Office, London.

Wu, J., Xia, J., and Horowitz, A. J. (2008). "Methodologies for estimating metered on-ramp vehicle queue length." *Proc., 87th Transportation Research Board Annual Meeting*, Washington, DC.