How to estimate queue length in real-time at signalized intersection is a long-standing problem. The problem gets even more difficult when signal links are congested. The traditional input–output approach for queue length estimation can only handle queues that are shorter than the distance between vehicle detector and intersection stop line, because cumulative vehicle count for arrival traffic is not available once the detector is occupied by the queue. In this paper, instead of counting arrival traffic flow in the current signal cycle, we solve the problem of measuring intersection queue length by exploiting the queue discharge process in the immediate past cycle. Using high-resolution "event-based" traffic signal data, and applying Lighthill–Whitham–Richards (LWR) shockwave theory, we are able to identify traffic state changes that distinguish queue discharge flow from upstream arrival traffic. Therefore, our approach can estimate time-dependent queue length even when the signal links are congested with long queues. Variations of the queue length estimation model are also presented when "event-based" data is not available. Our models are evaluated by comparing the estimated maximum queue length with the ground truth data observed from the field. Evaluation results demonstrate that the proposed models can estimate long queues with satisfactory accuracy. Limitations of the proposed model are also discussed in the paper.

© 2009 Elsevier Ltd. All rights reserved.
2. Shock wave analysis and “break points identification

2.1. Shock wave analysis

The traditional Lighthill–Whitham–Richards (LWR) traffic flow model hypothesizes that flow is a function of density at any point of the road. A shockwave is defined as “the motion (or propagation) of an abrupt change (discontinuity) in concentration” (Stephanopoulos and Michalopoulos, 1979). Traffic shockwave theory is derived from LWR model when applying the method of characteristics to analytically solve the partial differential equation (PDE) in the model. Basically, when characteristic curves (along which, the density is constant) interact, a shockwave is formed and wave velocity can be determined by following equation (Eq. (1))

\[ u_w = \frac{\Delta q}{\Delta k} = \frac{q_2 - q_1}{k_2 - k_1} = k_2 u_2 - k_1 u_1 \]

where \( q_1, k_1, u_1 \) are the flow, density and velocity of the upstream region and \( q_2, k_2, u_2 \) are the flow, density and velocity of the downstream region. Traffic shockwave can be also illustrated by using the fundamental diagram \((q-k)\) curve. The tangent of the chord drawn between any two points on the \( q-k \) curve defines the shockwave speed, as shown in Fig. 1 (we will explain \( \psi_1, \psi_2, \) and \( \psi_3 \) in the following).
At signalized intersections, multiple shock waves are generated due to the stop-and-go traffic caused by signal changes. For better explanation simply assume that queue has been fully discharged during the last green phase. In the following red interval, vehicles are forced to stop, which creates different flow and density conditions between the arrival and the stopped traffic. Such interruption of traffic flow, as indicated in Fig. 2a, forms a queuing shockwave $v_1$ in Fig. 1 moving upstream of the intersection with velocity

$$v_1 = \frac{0 - q_a^n}{k_j - k_a^n}$$

where $0$ and $k_j$ represent the jammed flow and density; and $q_a^n$ and $k_a^n$ are the average arrival flow rates and density during the $n$th cycle (subscript $a$ means "arrival traffic"). In Fig. 2, $T^g_n$ and $T^r_n$ indicate the end time and the start time of the effective green, respectively, during the $n$th cycle.

The queuing shockwave $v_1$ keeps propagating upstream. At the beginning of the effective green ($T^g_n$ in Fig. 2b), vehicles begin to discharge at saturation flow rate (assume there is no congestion downstream) forming the second shock wave which is defined as discharge shockwave $v_2$ in Fig. 2 at the stop line moving upstream with speed

$$v_2 = \frac{q_m - 0}{k_m - k_j}$$

where $q_m$ and $k_m$ are the capacity or saturation flow and density.

The discharge shockwave $v_2$ usually has higher speed than $v_1$, so the two waves will meet at time $T^{\text{max}}_n$ (in Fig. 2c), which is the time that this approach has the maximum queue length. As soon as the two shock waves meet, a third one (defined as a departure shockwave, $v_3$ in Fig. 1) is generated propagating toward the stop line with speed

$$v_3 = \frac{q_m - q_a^n}{k_m - k_a^n}$$

At the end of this cycle, i.e., the start of the red phase of the next cycle, if the queue cannot be fully discharged, a residual queue is formed, which builds the fourth shock wave $v_4$ (Fig. 2d) at stop line moving upstream with speed

$$v_4 = \frac{0 - q_m}{k_j - k_m}$$

Shock wave $v_4$ describes the queue compression process. Waves $v_3$ and $v_4$ have inverse directions therefore they meet at time $T_{\text{min}}^n$ (Fig. 2e), which is the time the approach has minimum queue length, i.e., the residual queue. As soon as the two waves meet the fifth wave $v_5$, which is the new queuing wave of the $(n+1)$th cycle, is formed moving upstream with similar speed as the shock wave $v_1$.

$$v_5 = \frac{0 - q_a^{n+1}}{k_j - k_a^{n+1}}$$

Similar process is repeated in the following cycles as indicated from Fig. 2a–f.

It is clear that the tail end of the queue follows the trajectory of the shock wave $v_1$ and $v_3$. However, the queue dynamics can be analytically calculated only if accurate vehicle arrivals are known. As we mentioned before, in the real world, once vehicular queue spills back to detectors, traffic arrival information at the detector location in the current cycle is not available until queue starts to discharge. If the event-based traffic signal data can be stored and archived, then such data can be utilized to estimate queue length in the immediate past cycle. The shockwave motion described above demonstrates the repetitive queuing and discharging process at a signalized intersection. The queues begin to accumulate at the start of...
red interval and discharge at the beginning of green phase. Sometime after the green start, maximum queue length is achieved; and shortly after the red start, minimum queue length (i.e., residual queue) is formed. Such repeated process gives us some hints that if some crucial “break points” can be identified utilizing detector data, then the queuing and discharging

Fig. 2. Shock wave propagation.
process can be recovered for the past cycles. As indicated in Fig. 3, most of waves propagate through detector line (if there exist waves that do not cross the detector line, then this link either has short queues or is oversaturated, which will be discussed later). The points (A, B, and C in Fig. 3) represent the times that traffic flow changes at the detector location. These curial points, i.e., break points, if recognized, can be used to recover the queuing process in this cycle. In particular, we are interested to estimate the maximum queue length in a cycle, i.e., point H in Fig. 3. We should note that the “fill-up time” used in Muck (2002) refers to the time difference between the start of red \( T_g \) and \( T_A \). As will be introduced in the next, high resolution traffic signal data can be utilized to identify these points.

2.2. Description of SMART–SIGNAL data

It is necessary to briefly introduce the high resolution data before we discuss how to identify break points. The data utilized in this research is collected by the SMART–SIGNAL system, which has successfully been installed on two major arterials in the Twin Cities area. SMART–SIGNAL can continuously collect and archive high resolution event-based data including vehicle events and signal events (Fig. 4). Event data provide start and end times of each vehicle-detector actuation “event” and every signal phase change “event”. Therefore, the time difference between the start and the end of a signal event presents the phase interval; and the time interval between the start and the end of a vehicle actuation event is the detector occupancy time, which is also the time that a vehicle passes the detector. The occupancy time can be used to calculate percentage occupancy; and is an indicator of the travel speed; as by assuming an effective vehicle length, the travel speed can be estimated. In addition, the time interval between the end of a vehicle actuation event and the start of next vehicle actuation event (from the same detector) is the vehicle gap, which is the time interval between two consecutive vehicles crossing the detector.

2.3. Break points identification

As mentioned in Section 2.1, “break points” A, B, and C represent the time instants that traffic condition changes within a cycle (here we define the cycle start is the effective red start and the cycle end is the effective green end). In detail, the time that point A appears \( (T_A) \) is the moment that the queuing shock wave \( v_1 \) propagates backward to the location of the loop detector. Between \( T_g^n \) (the end of green in the \( n \)th cycle) and \( T_A \), the vehicles pass the loop detector with the traffic state \( (q_0^n, k_0^n) \); while between \( T_A \) and \( T_{mA} \) (the time of maximum queue achieved), no vehicle can pass the loop detector because of the jam traffic condition \((0, k_j)\). Point A can be used to judge whether there is a long queue or not, because if point A does not exist, which means that the queuing shockwave does not propagate to the detector, then the queue length is less than the distance between the stop line and the advance detector. Point A is not difficult to identify; as after \( T_{mA} \), the detector is occupied for a relatively long time, so the value of the detector occupancy time is relative large. A threshold value is necessary for practical application. In this study, based on our observation, 3 s is a large enough number to check whether point A exists. If the detector occupancy time is larger than 3 s after \( T_{mA} \), the intersection has long queue; and vice versa. Fig. 5a demonstrates “event-based” detector occupancy time in a sample signal cycle. As indicated in the figure, after queue spills back
to the advance detector, the occupancy time is significantly larger than 3 s (about 45 s in this particular cycle). We should point out that second-by-second percentage occupancy data can also be utilized to identify point A, i.e., the occupancy value is kept at 100% for more than 3 s.

Point B indicates the time ($T_{B}$) that the discharge shockwave passes the detector. Between effective green start ($T_{n}^g$) and, $T_{B}$ the traffic state over the detector is ($0, k_{j}$); after $T_{B}$, vehicles are discharged at saturation flow rate and traffic state changes to ($q_{m}, k_{m}$). It is also not difficult to identify point B using high resolution data. After the green starts and before $T_{B}$, traffic volume is zero, and detector occupancy time is high (larger than 3 s) or second-by-second percentage occupancy continues to be 100% for at least 3 s. After $T_{B}$, queued vehicles begin to discharge over the detector, therefore both detector occupancy time and time gap between consecutive vehicles drop. As indicated in Fig. 5a, the detector occupancy time drops significantly after $T_{B}$.

The most important break point is point C. As will be introduced in the next section, the time of point C ($T_{C}$), combined with the discharge shock wave $v_{B}$, will be utilized to estimate the maximum queue length and re-construct queue forming and discharging process. Point C indicates the time ($T_{C}$) when the rear end of queue passes the detector. The time duration

---

**Fig. 4.** Sample data collected at the intersection.

**Fig. 5.** (a) Detector occupancy profile in a cycle; (b) time gap between consecutive vehicles in a cycle.
between $T_g$ and $T_c$ is closely related to the estimation of maximum queue length. As introduced before, wave $v_3$ is the interface between saturation traffic state $(q_m, k_m)$ and the arrival traffic state $(q_n^a, k_n^a)$. Therefore, before point C appears, vehicles discharge at the saturation flow rate at the location of loop detector, i.e., the traffic state is $(q_m, k_m)$. After the wave propagates to the detector location, the traffic condition becomes to $(q_n^a, k_n^a)$, i.e., the discharge rate at the loop detector location is less than saturation flow. A threshold should be selected to identify the two different traffic states $(q_m, k_m)$ and $(q_n^a, k_n^a)$. Based on our observation, the time gap between two consecutive vehicles is sensitive to traffic state change. As indicated in Fig. 5b, traffic is separated into two states by break point C. Before $T_c$, the time gaps between vehicles are small (less than 2.5 s) and the variance is small. It means that most of vehicles are discharged at saturation flow rate. But after $T_c$, the vehicle gaps become much bigger and the variance is significantly increased. More importantly, there usually exists a time lag between the saturated queue discharge flow and newly arrival traffic, as shown in Fig. 5b.

Ideally, statistical analysis of vehicle gap data (for example, comparing the means and variances of time gaps of two vehicle groups) is necessary in order to recognize traffic state change. However, for simplicity and practicality, a threshold value can be estimated using detector data. Therefore, in this research, wave

$$v_3$$

is identified. Considering the variation of time gaps, using a single value to separate traffic states may bring large error. In our implementation, if the time gap is between 2.5 s and 3 s, the system will continue searching the second and third points with time gaps over 2.5 s to make sure that the traffic state is really changed. It should be also noted that if only second-by-second data is available, the method described above is still feasible. As we know that larger than 2.5 s gap means 0% occupancy for at least two consecutive seconds.

Two other interesting cases are demonstrated in Fig. 6, one is that the break point C cannot be found; and the other is that vehicles arrive as a platoon. In Fig. 6a, during the green phase, the traffic pattern does not change and vehicles keep discharging at saturation flow rate. This case could be caused by two reasons: (1) vehicle queue cannot be discharged completely and this approach at this cycle is possibly oversaturated; (2) vehicle platoon arrives within a small time lag so that a large vehicle gap cannot be found. The other case, as indicated in Fig. 6b, is that although the break point is easy to find, the traffic pattern before and after break point, however, is similar. The explanation of this situation is that after the queue has been discharged, a platoon discharged from upstream intersection arrives. This situation, sometimes, will bring error to our queue estimation model, which will be discussed in the later sections.

3. Queue estimation models

The models proposed here are to utilize the break points (A, B and C) identified in the last section using high resolution data. As indicated in Fig. 3, the crucial part is how to estimate the coordinate of point H, the maximum queue, in both spatial (i.e., queue length) and temporal ($T_{max}$) dimensions. Point H is the intersection point of three waves and its coordinate can be decided by any two of them. As mentioned above, wave $v_1$ is a queuing wave, which highly depends on traffic arrival. Although wave $v_3$ is also related to arrival flow (Eq. (4)), it is the arrival traffic after queue discharge so that the traffic state can be estimated using detector data. Therefore, in this research, wave $v_2$, which has a constant velocity (as shown in Eq. (3) if we assume saturation flow rate $(q_m)$, saturation density $(k_m)$, and jam density $(k_r)$, are known a priori), and wave $v_3$ are utilized to identify the coordinate of point H, i.e., the location and time when maximum queue is reached. One basic model and two expansions, which are specially designed for some practical applications, are discussed in this section.

3.1. Basic model

To estimate shockwave speeds, $v_2$ and $v_3$, we need to "mine" more information from high resolution event-based data from one single loop detector. Shockwave speed $v_2$ can be estimated using the distance between advance detector and stopbar ($L_d$) and the time difference between the green start ($T_g$) and the time discharge wave reaching advance detector ($T_B$), i.e.,

$$v_2 = L_d/(T_g - T_B).$$

Shockwave speed $v_3$ can also estimated using Eq. (2) with assumed saturation flow rate $(q_m)$, saturation density $(k_m)$, and jam density $(k_r)$.

To estimate shockwave speed $v_3$, we will need to estimate two traffic states $(q_m, k_m)$ and $(q_n^a, k_n^a)$ and apply Eq. (4). This is feasible because event-based data record both the occupancy time and the gap between consecutive vehicles passing the loop detector. Occupancy time recorded by detectors can be used to estimate the individual speed by assuming an effective vehicle length; and the sum of occupancy time and vehicle gap, i.e., headway, can be utilized to estimate the average flow. Then density can be estimated by average flow and space mean speed. Eq. (7) is used to estimate the space mean speed ($u_i$), flow ($q$) and density ($k$)

$$\begin{align*}
&\left\{ \begin{array}{l}
  u_i = L_x/t_{oi} \\
  u_i = 1/(\frac{1}{n} \sum_{i=1}^{n} t_{oi}) \\
  q = 1/(\frac{1}{n} \sum_{i=1}^{n} h_i) = 1/(\frac{1}{n} \sum_{i=1}^{n} (t_{oi} + t_{gi})) \\
  k = q/u_i
\end{array} \right.
\end{align*}$$

(7)
where, $t_{o,i}$ and $t_{g,i}$ are the detector occupancy time and time gap of vehicle $i$, respectively; $u_i, h_i$ are the speed and headway of vehicle $i$, respectively; $L_e$ is the effective vehicle length, i.e., the sum of average vehicle length and detector length, which may require calibration; and $n$ is the number of vehicles which has been identified with same traffic state.

Fig. 6. Two other cases – (a) oversaturation; (b) platoon arrival.
As we discussed before, the traffic state is \((q_m, k_m)\) between TB and TC, and \((q_n, k_n)\) after TC (and before \(T_{n+1}\)). Eq. (7) can then be applied to estimate these two conditions. Note that in the algorithm, we use observed data to estimate traffic state \((q_m, k_m)\) instead of assuming the constant values; this is because the capacity or saturation flow rate will decrease if the traffic is affected by the downstream intersections. The velocity of wave \(v_3\) is calculated based on the two estimated traffic conditions.

Using estimated \(m_3\) and \(v_2\), the maximum queue length \(L_{n_{\text{max}}}\) and time \(T_{n_{\text{max}}}\) during \(n\)th cycle can be calculated:

\[
\begin{align*}
L_{n_{\text{max}}} &= L_d + \left(T_C - T_B\right)\left(\frac{1}{v_2} + \frac{1}{v_3}\right) \\
T_{n_{\text{max}}} &= T_B + \left(L_{n_{\text{max}}} - L_d\right)/v_2
\end{align*}
\]  

(8)

where \(L_d\) is the distance from stop line to the loop detector.

The minimum queue length \(L_{n_{\text{min}}}\) and time \(T_{n_{\text{min}}}\) (if residual queue exists) during \(n\)th cycle can be also calculated:

\[
\begin{align*}
L_{n_{\text{min}}} &= \left(\frac{L_{n_{\text{max}}}}{v_3} + T_{n_{\text{max}}} - T_{n_{\text{max}}+1}\right)\left(\frac{1}{v_4} + \frac{1}{v_3}\right) \\
T_{n_{\text{max}}} &= T_{n_{\text{max}}+1} + L_{n_{\text{min}}}/v_4
\end{align*}
\]  

(9)

where \(v_4\) is the velocity of the shock wave \(v_4\), which has similar value as velocity of the shock wave \(v_2\) (Eq. (5)).

The dynamic queue discharging process after \(T_{n_{\text{max}}}\) (before \(T_{n_{\text{min}}}\)) is easy to be formulated as we know the wave speed \(v_3\); and the queuing process before TA has been recorded by loop detectors (as we can see in Fig. 7, the arrival may not be constant).

The problem now is to estimate the queue state between points A and H. Without any other information, we can assume a constant velocity for wave \(v_1\), which can be estimated by

\[
v_1 = (L_{n_{\text{max}}} - L_d)/(T_{n_{\text{max}}} - T_A)
\]  

(10)

Then the entire queue accumulation and discharge processes can be fully described.

3.2. Expansion I – using second-by-second detector data

In the basic model, event-based data is required in order to identify the shockwave speeds; such requirement may not be satisfied in the real world. If second-by-second detector data and signal phase data are available (such as the ACS-Lite system), the basic model for the calculation of maximum and minimum queue lengths can still be applied by identifying break points A, B and C and calculating traffic state variables using Eq. (7). Model Expansion I provides a simple alternative to calculate maximum and minimum queue length, without applying Eq. (7), as explained below.

As we know, among all the vehicles passing loop detectors before TC, most of them are part of the maximum queue \(L_{n_{\text{max}}}\), while a small portion of vehicles join the queue after the tail end of original maximum queue began to move. This small portion of vehicles theoretically has no contribution to the maximum queue length, but they are affected by the queue. There-
fore, an approximation can be made to treat all the vehicles passing the detector between the green start and $T_C$ as the queued vehicles. Because the number of vehicles can be easily obtained using second-by-second detector data, by assuming a constant jam density, the maximum queue length ($L_{n\text{max}}$) and the time ($T_{n\text{max}}$) as well as wave speed $v_0$ under such approximation can be directly estimated:

$$
L_{n\text{max}} = \frac{n}{k_j} + L_d
$$

$$
T_{n\text{max}} = T_r + \frac{L_{n\text{max}}}{v_2}
$$

$$
v_3 = \frac{v_{n\text{max}}}{v_2} - \frac{L_d}{(T_C - T_{n\text{max}})}
$$

where $n$ is the number of vehicles passing detector between $T_g$ and $T_C$.

By using a constant shock wave speed $v_4$, Eq. (9) can then be applied to estimate the minimum queue length ($L_{n\text{min}}$) and the time ($T_{n\text{min}}$) as well as wave speed $v_0'$.

Essentially, Expansion I is a simplified queue estimation model. As demonstrated in Fig. 8, point H represents the true maximum queue, while H' is an approximation. Theoretically, the expanded model I overestimates the maximum queue length as it includes some portion of vehicles which do not belong to the stopped vehicle queue.

3.3. Expansion II – dealing with wired-together detectors

The expanded model II is specifically designed to deal with a case that detectors in different lanes (in a same approach) are wired together, or a single detector cover multiple lanes. In this case, vehicles traveling at different lanes will be counted only once if they pass the detector at the same time. This detector configuration is common for many signalized intersections in Minnesota. Under such design, although the break points can still be identified by using second-by-second detector data, the vehicle counts are not accurate any more (which means that the Basic Model and Expansion I will not work). An approximation is made here to deal with this problem. This approximation is based on the assumption that after the tail end of maximum queue begins to move, no additional vehicles can join the queue. Then the shock wave $v_{n\text{max}}'$ actually is the trajectory of the last vehicle. By assuming a priori known constant acceleration speed (3.5 feet/s, for example), the trajectory of the lasted stopped vehicle can be analytically derived.

Depending on whether the last vehicle in the queue reaches maximum speed at point C, the time interval between $T_{n\text{max}}'$ and $T_C$ (in Fig. 9) can be estimated using following equation:

$$
\begin{cases}
\frac{1}{2} a\left(\frac{u_f}{a}\right)^2 + u_f\left(t - \frac{u_f}{a}\right) = v_2(T_C - T_B - t) & \text{if } t \geq \frac{u_f}{a} \\
\frac{1}{2}at^2 = v_2(T_C - t) & \text{if } t < \frac{u_f}{a}
\end{cases}
$$

where $u_f$ and $a$ are free flow speed and acceleration speed, respectively.
Then the maximum queue length \( (L_{\text{max}}') \) and the time \( (T_{\text{max}}') \) can be estimated by:

\[
\begin{align*}
L_{\text{max}}' &= L_d + v_2(T_C - T_B - t) \\
T_{\text{max}}' &= T_C - t
\end{align*}
\]

(13)

Similar formulations are used to estimate the minimum queue length \( (L_{\text{min}}') \) and the time \( (T_{\text{min}}') \) as well as wave speed \( v' \). Clearly, the assumption made in this model may not be true in reality. So for some situation, similar as Expansion I, Expansion II overestimates the maximum queue. As indicated in Fig. 9, point H is the real maximum queue length, and estimated maximum queue length \( (H'') \) actually is an overestimation.

4. Implementation

4.1. Implementation procedure

Fig. 10 demonstrates the implementation procedure for the queue length estimation algorithms. To implement such algorithms, the first step is to check whether break point A exists. If point A cannot be found, it is a short queue case; and a simple input–output method can be applied to estimate queue size. This simple method records number of vehicles passing loop detector, assumes saturation flow rate discharging when signal turns green, and calculates the residual number of vehicles within the area between stop line and the location of the advance detectors. Then queue length can be estimated by simply assuming an effective vehicle length in jam traffic state. If point A exists, then it requires identifying point B and C. (actually, point B is not necessary to identify because the wave speed \( v_2 \) for most of situation is a constant value). If point C can be identified, depending on the resolution of detector data, different models are applied to estimate the long queue. If point C cannot be identified, as we discussed in the last section, the approach is under oversaturation. The queue length under oversaturated condition is difficult to estimate, but it is longer than or equal to the distance that occupied by the maximum dischargeable queue during green time.

4.2. Field evaluation results by the research team

The intersection of Trunk Highway 55 and Rhode Ave in Minnesota is selected as the testing site because of the frequent occurrences of extended queues on the west bound (From Glenwood Ave. to Rhode Ave.) during morning peak. Fig. 11a is a map of the intersection (including six coordinated intersections on Trunk Highway 55) and Fig. 11b shows the detector layout of the intersection. The data collected from detector #9 and #10 during two morning peak hours with fixed cycle-length (180 s) on June 11th, 2008 (Wednesday) are used to estimate the queue length; and the estimated values are compared with the ground truth data, which is recorded by a camera installed at the intersection (the maximum queue length for each cycle is manually extracted by watching the video). As event-based data are collected by detectors, all three models (basic,
Expansions I and II) can be tested. The results of the maximum queue length are presented in Fig. 12; and Table 1 shows the Mean Absolute Percentage Error (MAPE) which is calculated by

$$\text{MAPE} = \frac{1}{m} \sum_{i=1}^{m} \left| \frac{\text{Observation} - \text{Estimation}}{\text{Observation}} \right| \times 100\%$$

where $m$ is the total sample size.
As presented in Fig. 12, for both left lane and right lane, the basic model can successfully estimate the maximum queue length; and the average MAPE is around 7% (Table 1). However, Expansion I and Expansion II, as we mentioned before, overestimate the maximum queue; and the average absolute percentage errors are about 14% and 20%. Fig. 13 visually presents the estimated queue dynamics. It is very clear that the proposed models successfully describe queue forming and discharging processes. In addition, Table 2 presents the time differences when the maximum queue length is reached ($T_{max}^n$), using the three proposed models and real-world observations. The absolute errors for three models are around 5 s.
4.3. Independent evaluation results by Alliant Engineering Inc.

A Minneapolis-based Transportation Consulting firm, Alliant Engineering Inc. also conducted an independent evaluation of the queue length estimation algorithm. To observe the queue length, Alliant sent observers to the field (the Rhode Island intersection) during morning peak (7:00–9:00 a.m.) on three randomly selected days in 2008: July 23rd, October 29th, and December 10th. These observers manually count the vehicles as they enter the queue (they were instructed to count a stopped vehicle as one that was traveling at less than 5 miles per hour) and record the time when queue is maximal.

<table>
<thead>
<tr>
<th>Absolute errors of time of maximum queue length</th>
<th>Basic model (s)</th>
<th>Expansion I (s)</th>
<th>Expansion II (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left lane (Detector #10)</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Right lane (Detector #9)</td>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2

Figure 14. Comparison of maximum queue length.
Fig. 14 compares the time and length of maximum queue estimated by the basic model and the observations (only right lane results are presented here). As indicated in the figure, the proposed model tracks the trend of cycle-based queue dynamics successfully. The MAPE is relatively high (14.93%, on average) compared with our own evaluation due to two possible reasons: (1) The collected queue size data (i.e., the number of queued vehicles) are converted to queue length by simply multiplying 25 feet (the assumed constant effective vehicle length). Variation of actual vehicle length in the field may bring in conversion errors. For example, it is no unusual that when the detector was occupied (equivalently queue length is 400 feet), there were only 10 or 12 vehicles within the area between stop bar and the detector line (equivalently queue length is 250 or 300 feet). (2) Engineering judgment was required during data collection to identify whether vehicles join queue. This may also involve some errors. Overall, the proposed model performed very well in tracking cycle-based queue length changes for all three testing scenarios (see Table 3).

### 5. Discussion

Although we are satisfied with the performance of the proposed models, it is necessary to discuss model limitations.

1. **Break point C identification**. The identification error occurs when arrival traffic is at saturation flow rate and time lag between the arrival flow and queue discharge is equal to saturation headway. For example, as indicated in Fig. 15, a platoon discharged from upstream intersection arrives right after the end of queue starts to discharge. In such situation, the break point C cannot be identified, potentially leading to estimation errors. We argue, however, the likelihood for the occurrence of such situations is very small.

2. **Oversaturation**. The proposed models will work properly when break point C can be correctly identified. If point C cannot be identified, this approach of the intersection is considered as possibly “oversaturated” (excluding the cases indicated above). In such situation, the signal link is experiencing saturated traffic flow, i.e., density is $k_{\text{max}}$. It is also necessary to point out here that in this paper we do not consider the case with extreme congestion, i.e., the link has totally been blocked by the downstream queue, so that point C cannot be identified. Current model has limitation to deal with such case as the shockwave profile is different under this condition.

3. **Accuracy comparison among models**. In the three proposed models, theoretically, the basic model is more accurate than the other two. This is also supported by the testing results. However, since the basic model uses an additional parameter (pre-determined effective vehicle length) to estimate individual vehicle speed, it may bring errors to the queue estimation if the effective vehicle length is not accurate. In situations such as frequent platoon arrivals, since the trajectory of the last queued vehicle is identical to the shock wave trajectory, Expansion I & II are more suitable.

![Fig. 15. Immediate platoon arrival after queue discharge.](image-url)
Detector errors. We need to point out that the proposed models do not consider the impact of detector errors, such as miss-counting or over-counting. An error filtering method such as Kalman Filter, may improve the reliability of the proposed model. Such research is left for future study.

6. Concluding remarks

In this paper, we proposed an innovative approach to estimate the intersection queue length using existing detectors. A key methodological contribution of our approach is that it can estimate time-dependent queue length even when the signal links are congested with long queues. By applying LWR shockwave theory with the high-resolution traffic signal data, we are able to distinguish different traffic states at the intersection, so that queue length estimation under congested conditions becomes possible. Two expanded models are also provided for practical application to deal with the problem that if only second-by-second data is available and that if the detectors are wired together. Three models are evaluated by comparing the estimated maximum queue length with the ground truth data recorded by camera and human observers. The results indicate the basic model is more accurate but the other two are also acceptable. Limitations of the proposed models are also discussed. In the future study, we will deal with these limitations either by improving the models’ robustness or by utilizing additional data, such as the information from upstream intersections, to increase the estimation accuracy and reliability.

Acknowledgements

The authors would like to thank James Grube and Eric Drager of the Hennepin County Transportation Department, and Steve Misgen and Ron Christopherson of Minnesota Department of Transportation for their support for the field implementation of the SMART–SIGNAL system. The authors also appreciate the constructive comments offered by the anonymous reviewers of this paper.

References