Impact of ramp metering queue override on the capacity of an isolated freeway merge

Xingan (David) Kan, Ph.D.
(Corresponding Author)
Department of Civil and Environmental Engineering
Institute of Transportation Studies
109 McLaughlin Hall, University of California, Berkeley, CA 94720-1720
Phone: +1(408) 613-7638, Email: david.kan@berkeley.edu

Xiao-Yun Lu, Ph.D.
Partners for Advanced Transportation Technology (PATH)
Institute of Transportation Studies
Richmond Field Station Building 452, Richmond, CA 94804
Phone: +1(510) 665-3644, Email: xiaoyun.lu@berkeley.edu

Alexander Skabardonis, Ph.D.
Department of Civil and Environmental Engineering
Institute of Transportation Studies
109 McLaughlin Hall, University of California, Berkeley, CA 94720-1720
Phone: +1(510) 642-9166, Email: skabardonis@ce.berkeley.edu

For Presentation and Publication at the
Transportation Research Board
98th Annual Meeting
Washington, D.C.
January 13–17, 2019

August 1, 2018

Word Count: 3785 for text (4035 including tables)
Number of Tables: 1
ABSTRACT

Freeway ramp metering is an effective control strategy to preserve freeway capacity, reduce freeway delays, and improve travel time reliability. The ramp metering operation may cause on-ramp queue spillback that interferes with the adjacent surface streets. In such situations, most ramp metering systems employ a “queue override” function, which temporarily suspends ramp metering to dissipate the on-ramp queues. A detailed field study was performed at a metered freeway on-ramp merge in San Jose, California using video recordings. The analysis of the collected field data over a two week period show that the freeway bottleneck discharge flow is reduced by 10% on the average when the queue override is activated. The paper provides suggestions for managing queue spillbacks at metered on-ramps.
INTRODUCTION

A bottleneck is defined as a point where the traffic demand exceeds the normal freeway capacity, resulting in formation of queues upstream of that location and free-flowing traffic downstream. The bottleneck is called “active” when traffic flow through the bottleneck is not affected by downstream restrictions (spillback from downstream bottlenecks). Recurrent bottlenecks occur on the same location and time periods of the day. Their behavior and characteristics are reproducible over many days. Typically the bottleneck remains active throughout the peak period(s). Traffic queues dissipate from the back as traffic demand drops below the available capacity. On the other hand, non-recurrent bottlenecks due to incidents generally have shorter duration, although some major incidents may last a long time. Non-recurrent bottlenecks are non-reproducible since incidents are random events and may occur anywhere in the freeway system. Furthermore, traffic queues dissipate from the front following the incident removal, i.e. when the normal capacity is restored.

There are various bottleneck types depending on their location and causes. At freeway merging areas, vehicles entering from the on-ramp trigger traffic breakdown forming a bottleneck in the main lanes shortly downstream of the merge point. The capacity of a freeway bottleneck is defined as the maximum sustained flow it discharges under prevailing traffic and roadway conditions, provided that the freeway’s outflow is not impeded by exogenous restrictions such as queue spillback from further downstream. High entry volumes at on-ramps cause flow breakdown that reduces the discharge flow of the freeway bottleneck. This phenomenon is referred as “capacity drop” and has been well documented and examined in many empirical studies (1-9). These empirical studies suggest that capacity drop typically entails a 5% to 15% reduction in the bottleneck discharge flow. In addition, many have proposed mathematical models to explain capacity drop (10-13). The latest edition of the Highway Capacity Manual recommends applying a 7% reduction to the freeway bottleneck capacity when there is significant merging traffic from an on-ramp, in order to account for capacity drop and breakdown near merging areas (14).

Freeway on-ramp metering has been extensively used as a traffic control strategy to regulate the entry of the on-ramp vehicles in order to prevent congestion and preserve the freeway capacity, thus avoiding the capacity drop. The effectiveness of ramp metering has been demonstrated in several field studies (5, 15, 16), as well as simulation studies (17-20). Additional benefits of ramp metering include accident reduction, improved freeway travel time, and better travel time reliability (21).

The ramp metering operation may create long queues at the on-ramps that may exceed the queue storage and interfere with the operation of the adjacent surface street network. This is a common occurrence on California freeways because most of the on-ramps do not provide sufficient queue storage (22). Most of the operational ramp metering systems employ a “queue override” feature that is intended to prevent the on-ramp queue from obstructing traffic conditions along the adjacent surface streets (23, 24). The override is triggered whenever a sensor placed at the entrance of the on-ramp detects a potential queue spillover of the on-ramp vehicles on the adjacent surface streets. This clears the on-ramp queue by temporarily turning off ramp metering. Unfortunately, this approach may reduce the effectiveness of the employed ramp metering systems during the time of the highest traffic demand, when the ramp metering is most needed. Currently, the Highway Capacity Manual does not include analysis procedures to account for both impacts of ramp metering on surface streets, the impacts of queue spillback, as well as the effect of queue override on freeway bottleneck capacity. This has been recognized as
a significant research need and is expected to be addressed in an upcoming NCHRP research project (25).

The objective of the study described in this paper is to provide empirical evidence on the queue override on the capacity of freeway merge bottlenecks, and suggest possible approaches to managed on-ramp queue spillback when freeway ramp metering is in operation. The research is part of a larger research project on freeway arterial coordination.

The rest of the paper is organized as follows: The next section provides an overview of the current state of queue override and its relevant literature. The following section summarizes the locations and procedures used to collect empirical data. The next section analyzes the observations made based on the field data. The final section summarizes the study findings and recommendations.

BACKGROUND: QUEUE OVERRIDE

Several transportation agencies that operate ramp metering on their freeways employ queue override as a part of their ramp metering algorithm to prevent queue spillback. In Seattle, Washington the ramp metering rate is increased based on how far the on-ramp queue has propagated upstream (26). In Nevada and Texas the ramp metering is temporarily suspended, operation known as queue flush (27, 28), once queue spillback is detected. In California (29) and North Carolina (30) the queue override algorithm increases the metering rate to the maximum allowable value (typically 900 veh/hour/lane) to alleviate on-ramp queue spillback. Furthermore, ramp metering in Minneapolis, Minnesota employs a combination of the queue override approaches adopted in California and Texas (31). There has been very limited empirical evidence that quantifies the impact of queue override on freeway bottleneck capacity. Chilukuri et al. (32) conducted an empirical study on the effect of queue flush for short time intervals of 30 to 75 seconds. However, there was no evidence in the study suggesting that the bottleneck was isolated from exogenous restrictions, and the queue override was the only cause of capacity drop. There is no other empirical study that quantifies the impact of queue override, especially for metering at maximum allowable rate, an approach employed in California and North Carolina.

DATA COLLECTION

This section describes the freeway bottleneck selected for this empirical study. Figure 1 shows the selected site, an isolated merge of northbound Interstate 680 in San Jose, California. The on-ramp consists of two lanes upstream of the ramp meter, and the two lanes merge into a single lane before reaching the freeway mainline. The ramp meter restricts the flow of on-ramp merging traffic and ensures smooth merging operation of the two on-ramp lanes by alternating the green times assigned to each on-ramp lane. The metering system operates under the local traffic responsive demand-capacity approach. The metering rates are assigned based on thresholds of freeway mainline occupancies immediately upstream of the merging area. The maximum sustained flow of this bottleneck is typically observed during the morning peak (7:00 AM – 9:00 AM).
Figure 1. Study site: northbound Interstate 680 in San Jose, California.

The queue override algorithm of this site, typically activated from 7:30 AM to 8:00 AM, increases the ramp metering rate by 100 veh/hour/lane as soon as queue spillover is detected at the entrance of the on-ramp, and by another 100 veh/hour/lane if queue spillover continued in the next 30 second cycle, until it reaches the maximum rate of 900 veh/hour/lane. Video cameras were placed upstream and downstream of the McKee Rd. on-ramp merge during the study periods of May 9, 2016 to May 13, 2016 and May 16, 2016 to May 20, 2016, and the camera locations are shown in Figure 1. Video cameras instead of loop detectors were used to ensure high accuracy and better resolution. The camera placed upstream recorded all four mainline lanes, as well as the McKee Rd. on-ramp. The third camera was placed to ensure the absence of exogenous restrictions such as queue spillback from the bottlenecks further downstream. The records of the frequency and duration queue override activation were provided by Caltrans District 4, the agency that operates the freeway ramp meters at this site. Lastly, there were no major incidents or weather events during the selected study period.

Vehicle count at each location and each 30 second interval was extracted from the video data, and the bottleneck discharge rates during periods of active and inactive queue override, respectively, were compared.

OBSERVATIONS AND DATA ANALYSIS

Figure 2 shows the curves for cumulative vehicle count of the mainline lanes, obtained at all three cameras locations shown in figure 1, vs time, $t$. The curves were plotted to display virtual departures as a function of time at location 3 in figure 1 (33). The vertical displacement between the curves is the excess accumulation on the freeway segment of interest due to the limited capacity, and the area between the curves indicate the total delay of the freeway system. Figure 3 shows the curves for cumulative vehicle count of the on-ramp vs time, $t$. The data presented in the figures 2 and 3 were collected on Tuesday, May 10, 2016.
Figure 2. May 10, 2016: $Q(t)$ curves for locations 1 through 3.

Figure 3. May 10, 2016: $Q(t)$ curves for McKee Rd. on-ramp.
The vertical scales in figures 2 and 3 were modified by plotting on the oblique coordinate system, in order to make the excess accumulation (vertical displacement) clearly noticeable by visual inspection (34). \( O(t) \), the oblique coordinate transformation of the cumulative vehicle count, \( V(t) \), is described by the following:

\[
V(t) = v_0(t - t_0)
\]

where \( v_0 \) is the specified reference value of flow and \( t_0 \) is the specified reference value of initial time.

The \( O(t) \) curves shown in figure 2 reveal that the arrival rate at location 1 was relatively low and the freeway was free-flowing (all three curves overlap) from \( t = 7:00 \) to \( t = 7:13 \). Video data from location 1 also show that the observed on-ramp flow is relatively low at about 825 vph, as described by the \( O(t) \) curve in figure 3. This corresponds to the prescribed restrictive metering rate of 400 vph/lane for the period of \( t = 7:00 \) to \( t = 7:13 \). The variation in actual ramp flow can be attributed to variability in green times, driver behavior, etc.

Immediately after \( t = 7:13 \), the curves for locations 1 and 2 shown in figure 2 began to diverge as the freeway transitions from free-flow condition to queueing. At \( t = 7:15 \), the prescribed ramp meter rate increased from 400 vph/lane to 600 vph/lane, as indicated by the ramp flow of 1156 vph in figure 3. Despite the increase in on-ramp merging traffic, the bottleneck outflow remained high at 7524 vph during the initial period of queueing, as shown by the dashed lines.

Queueing continued at \( t = 7:30 \), when the prescribed ramp meter rate increased to 700 vph/lane (indicated by the ramp flow of 1356 vph in figure 3). Under the less restrictive ramp meter rate, the outflow of the bottleneck slightly increased to 7847 vph, shown in figure 2 by the dashed line.

However, the high outflow persisted only until \( t = 7:36 \), when sufficient on-ramp queue spillback prompted the activation of queue override based on records obtained from the Caltrans District 4. Queue override gradually increased the meter rate by 100 vph/lane every 30 second cycle until the meter rate reaches the maximum allowable value of 900 vph/lane. As indicated by the dashed line in figure 3, the on-ramp flow exceeded the expected 1400 vph after \( t = 7:36 \), at 1500 vph. The observed on-ramp flow was less than expected value of 1800 vph under the maximum meter rate because queues already formed at and near the merging area physically restricted the number of vehicles entering the freeway from the on-ramp. As shown in figure 2, queueing persisted after \( t = 7:36 \). The arrival rate remained high but the outflow of the bottleneck diminished, indicated by the downward trending \( O(t) \) curve at location 2.

Queue override continued but the on-ramp flow began to diminish from 1500 vph to 1352 vph at \( t = 7:52 \) because queue spillback occurred less frequently therefore queue override was not constantly activated. This explains the slight increase in the bottleneck outflow, indicated by the curve for location 2.

As shown in figure 3, queue override ended at \( t = 8:01 \) and the on-ramp flow returned to 1158 vph; this corresponds to the prescribed meter rate of 600 vph/lane for \( t = 8:00 \) to \( t = 8:18 \). Despite the relatively high on-ramp flow, the overall arrival rate at location 1 was relatively low, which led to free-flow conditions. The free-flow condition persisted after
when the on-ramp flow reduced to 814 vph due to the change in prescribed ramp meter rate; except for a brief period (t = 8:40:30 to t = 8:47:30) of surge demand that resulted in queuing and a high outflow of 7740 vph.

Further inspection of the $Q(t)$ curves for location 2 and location 3 reveals that the segment between these locations remained free-flowing for the entire study period (both curves always overlapped). Thus the bottleneck was isolated and located between location 1 and location 2. Furthermore, queue persisted during the period of queue override. Therefore, the observed reduction in the bottleneck outflow during $t = 7:36$ to $t = 8:01$ was not a result of a reduction in traffic demand nor the result of an exogenous downstream restriction but the result of queue override. According to figure 2, the bottleneck outflow during queue override diminished to an average of 6891 vph, a reduction of 12.18% in comparison with the bottleneck outflow immediately before queue override was activated.

Table 1 provides an overview of the observed freeway bottleneck capacities prior to and during queue override for the two week study period. There were slight variations in the percentages of capacity drop observed in different days. The observed capacities prior to and after the activation of queue override vary by the day of the week, for instance, the observed capacities on Tuesday May 10, 2016 and May 19, 2016 were higher than those of the other days. Furthermore, the duration of queue override and capacity drop was about 25 to 30 minutes on average, with the exception of a 15 minute duration on Tuesday May 17, 2016 and a 40 minute duration on Wednesday May 18, 2016. In addition to the day to day variation, the capacity drop can be slightly more severe during the first few minutes of queue override, for example, on Thursday May 19, 2016. Overall, the observations suggest that queue override diminishes the bottleneck outflow by an average of 10%.

### Table 1. Freeway bottleneck capacities during morning peaks.

<table>
<thead>
<tr>
<th>Freeway bottleneck outflow (vph)</th>
<th>Before queue override</th>
<th>After queue override</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Week 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 9, 2016 (Monday)</td>
<td>Not activated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 10, 2016 (Tuesday)</td>
<td>7847</td>
<td>6891</td>
<td>-12.81</td>
</tr>
<tr>
<td>May 11, 2016 (Wednesday)</td>
<td>6752</td>
<td>6058</td>
<td>-10.28</td>
</tr>
<tr>
<td>May 12, 2016 (Thursday)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 13, 2016 (Friday)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 16, 2016 (Monday)</td>
<td>Not activated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 17, 2016 (Tuesday)</td>
<td>7214</td>
<td>6672</td>
<td>-7.51</td>
</tr>
<tr>
<td>May 18, 2016 (Wednesday)</td>
<td>7109</td>
<td>6493</td>
<td>-8.67</td>
</tr>
<tr>
<td>May 19, 2016 (Thursday)</td>
<td>7532</td>
<td>6612</td>
<td>-12.21</td>
</tr>
<tr>
<td>May 20, 2016 (Friday)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>----</td>
<td>----</td>
<td>-10.30</td>
</tr>
</tbody>
</table>
Freeway ramp metering is an effective control strategy to preserve freeway capacity, reduce freeway delays, and improve travel time reliability. The ramp metering operation may cause on-ramp queue spillback that interferes with the adjacent surface streets. In such situations, most ramp metering systems employ a “queue override” function, which temporarily suspends ramp metering or relaxes the metering rate to dissipate the on-ramp queues. The activation of queue override reduces the effectiveness of ramp metering and increases freeway delay.

Currently, there are no empirical data on the impacts of queue override and freeway operating conditions. The Highway Capacity Manual does not have any analysis procedures to account for the impact of on-ramp queue spillback and queue override. There is a need to understand the impacts of queue override, and develop approaches for avoiding queue spillback.

A detailed field study was performed at a metered freeway on-ramp merge in San Jose, California using video recordings. The analysis of the collected field data show that the bottleneck discharge flow is reduced by 10% on the average when the queue override is activated. Several approaches have been proposed and implemented for managing queue spillback at metered freeway on-ramps. Extending the on-ramp to allow more queue storage is an effective geometric solution, but it cannot be implemented in most situations because of physical constraints and environmental concerns. A number of ramp metering algorithms include procedures to adjust on-line the metering rate to avoid queue spillback based on measurement of the on-ramp queue length. However, these approaches require extensive detector placement and very accurate real-time data, which make the algorithms hard to implement in most freeway control systems. Coordinated control of ramp meters and adjacent traffic signals may prevent spillback and a number of efforts are under way as part of the Integrated Corridor Management (ICM) initiative to manage facilities and systems comprised of freeways and arterial streets.

Independent signal control along the parallel arterials may lead to large platoons of vehicles entering the on-ramps with insufficient storage space creating queue spillback, which in turn blocks the arterial intersection reducing its carrying capacity. Kan et al. (37) recently developed a simple and readily implementable signal control strategy that adjusts the signal cycle length and green times at the adjacent intersection taking into consideration the on-ramp metering rate and storage queue. Simulation tests on a real world freeway-arterial corridor showed improvements in freeway and system-wide throughput and delay with modest delay increase for the arterial traffic. A field test of the strategy is planned for late 2018. The field test will provide additional empirical evidence on impacts of the implemented control of traffic performance that can be also used in Highway Capacity Manual methodologies development.

ACKNOWLEDGMENT

This research is supported by the State of California Department of Transportation (Caltrans) under Interagency Agreement No. 65A0518. We wish to thank the project technical panel for their guidance and support. The contents of this paper reflect the views of the authors, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California.

AUTHOR CONTRIBUTION

The authors confirm contribution to the paper as follows: study conception and design: Alexander Skabardonis, Xingan (David) Kan, Xiao-yun Lu; data collection: Xingan (David) Kan; analysis and interpretation of results: Xingan (David) Kan, Alexander Skabardonis, Xiao-yun Lu;
draft manuscript preparation: Xingan (David) Kan, Alexander Skabardonis. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES