

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

## **On-Ramp Metering and Commuter Delay: A Before and After Study**

**Kwangho Kim, Michael J. Cassidy**

**California PATH Research Report  
UCB-ITS-PRR-2010-10**

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

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification, or regulation.

Final Report for Task Order 6331

March 2010

ISSN 1055-1425



# **FINAL REPORT FOR PATH**

## **On-Ramp Metering and Commuter Delay: A Before and After Study**

By

Kwangho Kim and Michael J. Cassidy  
University of California  
Department of Civil and Environmental Engineering  
And the Institute of Transportation Studies  
109 McLaughlin Hall  
Berkeley, California 94720  
(510) 642-7702  
Cassidy@ce.berkeley.edu

December 2009



## **Executive Summary**

This report furnishes clear evidence that on-ramp metering can increase the output flow through a freeway, and by so doing diminish the total time that commuters collectively spend traveling on the freeway and its on-ramps. Empirical study was performed on a 6.3-mile stretch of northbound Interstate 5 in Sacramento. The stretch spans the interchanges of Pocket Road (to the south) to W street (See Figure 1). Traffic data, both from loop detectors and from videos, were collected during the morning rush periods over a period spanning several years. Data were initially collected in 2006 prior to the deployment of ramp meters at the site. Data were collected again in 2007 and 2008 after meters were installed on five on-ramps. (The meters operate using a control logic developed by Caltrans.) Finally, a metering logic was developed in response to certain traffic details observed at the site, and was tested there in spring and fall 2009. A number of interesting and useful findings resulted from all this, as described below.

As is typical of urban freeways, the study site contains a series of bottlenecks. A key one is caused by a horizontal curvature in the freeway alignment and is located toward the downstream end of the study area (See the detailed geometry in Figure 3). During morning rush periods, the output flows from this horizontal curve alternate between high rates (the desirable state of affairs) and significantly lower rates. Curiously, the periodicities of these alternating flows grew in the later portions of each rush; i.e. output flows early in a rush rose and fell every minute or two, while later in the rush the periods of high and low flows persisted for a longer durations. Low flow periods late in a rush typically persisted for about 6 minutes, and this led to longer-run average discharge flows that were significantly lower in the later portion of a rush, as compared with the early part of that rush.

We find that these variations in output flow are linked to the queue that forms (and dissipates) at the horizontal curve. In the early rush, the curve becomes a transient bottleneck; i.e. queues periodically form at the curve's entrance, persist for a short time and then disappear. The high (low) output flows coincide with the disappearance (presence) of the upstream transient queues.

Later in the rush, as demand gradually increased, queues that formed at the curve were no longer transient ones. Instead, these queues gradually expanded. Whenever their tails eventually reached the interchange of 43<sup>rd</sup> Ave 1-mile upstream, we observed an interesting phenomenon that has been termed the “pinch effect” (Kerner, 2002). Each queue’s arrival to that upstream interchange disrupted and severely constrained the flows there. (The queue’s presence at that interchange, mixed with the vehicular weaving maneuvers that occur there, is evidently a very damaging combination.) As a result, freeway flow at the 43<sup>rd</sup> Ave Interchange became so constrained (i.e. “pinched”) that a new, more restrictive bottleneck formed there. This new bottleneck starved the downstream freeway section of flow. The queue on that downstream section therefore gradually dissipated; i.e. the queue’s tail propagated forward through traffic as a kinematic wave. The elapsed time between a queue’s formation at the horizontal curve and the resulting pinch effect upstream was typically about 6 minutes, which explains the 6-minute drops in output flow periodically observed late in each rush.

Interestingly, we find that as the queue shrank downstream of the 43<sup>rd</sup> Ave Interchange (due to the pinch effect), drivers in that queue exhibited greater motivation to discharge through the horizontal curve; i.e. drivers tended to adopt short headways, such that discharge flows through the curve increased. This explains the recoveries in output flow that periodically occurred late in each rush. The high flows could not be sustained indefinitely, however. After some minutes, a new queue formed at the horizontal curve and the above process repeated. Thus, the pinch effect was typically observed multiple times during each morning rush.

Understanding this effect sheds light on what would constitute an effective strategy for metering the site’s on-ramps. The intent is to maximize output from the horizontal curve, since all else being equal this will minimize the total time spent traveling on the freeway and its on-ramps (e.g. see Cassidy, 2003). We thus formulated an on-ramp metering strategy that operates under the principles described below. To simplify the field tests of this strategy, its logic was applied only to the two on-ramps that reside nearest upstream of the horizontal curve (the on-ramps at Seamas and 43<sup>rd</sup> Ave); and all other metered ramps operated as per the Caltrans control logic.

1. In the early part of a rush, considerable effort was made to dissipate the curved bottleneck's transient queues as soon as they form. This entails the deployment of a more restrictive metering rate at the closest upstream ramp (Seamas). Whenever the queue dissipated as a result, a very relaxed metering rate was deployed at Seamas. The on-ramp immediately upstream (43<sup>rd</sup> Ave) was metered at a very relaxed rate early in a rush because the queue rarely grew long during these (early) periods.
2. Whenever the curved bottleneck's queue grew long in the late part of a rush, we re-doubled our efforts to dissipate it: metering became even more restrictive at Seamas and restrictive metering was implemented at 43<sup>rd</sup> as well.
3. Whenever the pinch effect arose, it took its course. During these times, restrictive metering was maintained at both the Seamas and the 43<sup>rd</sup> on-ramps.
4. Whenever a queue dissipated at the curved bottleneck thanks to the pinch effect, relaxed metering was deployed at both Seamas and 43<sup>rd</sup> until new queue formed at the horizontal curve on the freeway. This relaxation served to i) flush the ramp queues that typically grew long during the period of restrictive metering; and ii) saturate the curve with flow. This second point limits the likelihood of over-controlling the freeway.
5. Steps 2~ 4 above were repeated through the late portions of each rush.

Occupancies measured by the site's freeway (mainline) detectors were used to monitor queue expansion in real time. The field tests indicate that the above scheme is effective. Under the scheme, we observed that the morning rush period's long-run average output flow from the horizontal curve increased by about 5.5%, as compared with that measured in the absence of metering (in 2006). Coarse analysis indicates that for present-day demand, an outflow increase of this magnitude would diminish morning delay by roughly 30%. (This would translate to a reduction of about 430 vehicle-hours each morning rush and accounts for the delays collectively encountered on the freeway stretch and its on-ramps.)

We concede that the effectiveness of our metering scheme was aided by California's economic downturn, which generated lower demand at the site in 2009 as compared with

2006 and thus made it easier for our scheme to combat freeway queueing<sup>1</sup>. This fact, however, does not diminish the significance of our finding: the evidence shows that ameliorating a freeway queue (by some means) can increase bottleneck output flow and thereby diminish commuter delay. Thus the work demonstrates the value of devising freeway traffic control measures that are commensurate with whatever demand level prevails. In the case of high demand (e.g. due to a strong economy), one might develop metering schemes that combat freeway queues by coordinating control over numerous on-ramps, and not just at two ramps as was done in our simple field experiments.

Finally, we find that the metering logic presently used by Caltrans is also effective in combating the freeway queue and thus generates higher output flows from the horizontal curve. We note as an aside that the evidence shows that the Caltrans logic is not quite as effective as the one we have proposed. This finding is not particularly relevant from our point of view, however. The objective of our study has not been to compare one metering algorithm against another, but rather to explore the potential for ramp metering to diminish system-wide delay. The study's success in this regard owes much to the extraordinarily high level of collaboration and assistance given us from Caltrans personnel at HQ and district 3. We are much indebted.

---

<sup>1</sup> During a morning rush, average demand for the curve section of freeway was more than 400 vph lower in 2009 than in 2006.



## **TABLE OF CONTENTS**

1. Introduction	1
2. Literature review	2
3. Observed traffic phenomenon at the site	6
4. Ramp metering field tests	16
5. Conclusions	25
Acknowledgements	27
References	



# 1. Introduction

This report describes a study of the 6.3-mile stretch of northbound I-5 site that is highlighted with bold dashed line in Figure 1. The study began with detailed empirical analyses of traffic on the site during morning rush periods. The data were collected both from the sites' loop detectors and from videos; and analyses were performed for numerous mornings in 2006 through 2008. Findings show that serial bottlenecks on the site interact and exert significant influence on traffic delays and queuing; that one particular bottleneck formed by a horizontal curvature in the freeway alignment plays a key role in this; and that these damaging influences occurred primarily along the freeway stretch demarcated with a lightly-drawn box in Figure 1.

An on-ramp metering logic was developed in light of these observations. Field tests of this logic were conducted over multiple morning rush periods in 2009. Findings show that ramp metering can affect traffic flow through the sites' serial bottlenecks in favorable ways that increase output flows and thereby reduce the delays collectively incurred by commuters on the freeway site and its on-ramps.



Figure 1 Study Site, Northbound I-5, Sacramento, California

A literature review of subjects relevant to the present study is furnished in the following section of this report. Observed details of traffic flow on the site, including the mechanisms that affect its output flows, are presented in section 3. Description of the proposed ramp metering scheme, and the outcomes of the field tests of this scheme are given in section 4. Conclusions are summarized in Section 5.

## **2. Literature review**

This section summarizes the literature relevant to two research areas: (i) freeway traffic flow phenomenon through serial bottlenecks, and (ii) on-ramp metering experiments to increase discharge flow from freeway bottlenecks. Although kinematic wave theory has often been used to describe real-world traffic by approximating it as fluid, some studies have reported certain traffic details that are not described well by kinematic theory (Sec 2.1). And although there have been a number of field experiments attempting to increase the discharge flow from freeway bottlenecks, many of these studies have not furnished definitive results (Sec 2.2).

### **2.1 Freeway traffic flow phenomenon in serial bottlenecks**

Lighthill and Whitham (1955) and Richards (1956) were the first to postulate that traffic propagates as kinematic waves. Newell (1993) simplified the application of kinematic theory by assuming that the relation between traffic flow and density (the Fundamental Diagram) is triangular in shape; and applied the theory using curves of cumulative vehicle counts at fixed freeway locations. The theory has been used to provide realistic descriptions of some key phenomena at freeway bottlenecks (e.g. Daganzo et al, 1999, Bertini, 2005).

Additionally, researchers have observed other traffic details which depart somewhat from kinematic wave theory. Some of the literature (Banks, 1991, Hall et al, 1991, Persaud, et al, 1998, etc) report that substantial discharge flow reductions occur at an

active bottleneck<sup>2</sup> where queues form upstream. This so called capacity drop phenomenon was observed recurrently whenever the density upstream of a bottleneck exceeded a certain critical value (Cassidy and Rudjanakanoknad, 2005; Chung et al, 2007).

The cause of capacity drop has been controversial among researchers. Kerner (1999) reported that capacity drop can be caused by random local perturbation unrelated to a bottleneck and then remain self-maintained, although he also conceded that bottlenecks are the most frequent reasons for this drop. Daganzo, et al (1999) argued that local perturbations, if not related to any bottleneck, dissipate after temporarily slowing down vehicles upstream (i.e. local perturbation cannot be self-maintained). In line with this argument, Daganzo (2002) and Laval (2004) proposed theories to explain capacity drop of an on-ramp merge bottleneck by means of lane-changing and driver car-following behavior.

Other literature reported that a freeway bottleneck's discharge flow can fluctuate after the queue forms upstream. Daganzo et al (1999) noted that substantial reductions in bottleneck discharge flow can be followed by recoveries in discharge flow, although recovered flows were sustained only for short periods of several minutes. In addition, Rudjanakanoknad (2005) reported that sequential surges and drops in queue discharge flow are more prominent at on-ramp merge bottlenecks with few (e.g. 2 or 3) freeway lanes.

A possible explanation of the time varying queue discharge flow can be found in a theory of traffic flow that Kerner and Rehborn (1996) proposed. This theory was used to explain an unusual mechanism called the 'pinch effect' (Kerner, 2002). The effect is described as below.

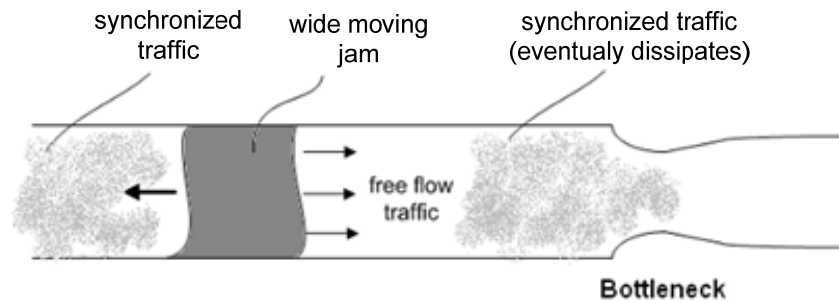
Kerner and Rehborn (1996) classified congested traffic phases either as (i) synchronized flow or (ii) wide moving jams. Synchronized flow is the congested traffic phase in which the average speeds of queued vehicles are nearly the same across lanes. In this state, vehicle speed is lower than free flow speed, and the flow can approach capacity. A wide moving jam is a localized traffic phase which is confined by two fronts that move

---

<sup>2</sup> An active bottleneck is characterized by queue discharge that is not affected by downstream traffic conditions (Daganzo, 1997). Unless these conditions are present, one cannot be certain that measured flows correspond to bottleneck capacity.

in the upstream direction. Inside a wide moving jam, the vehicle speed and flow are very low, even zero. The formation of a wide moving jam can dissipate a bottleneck's queue as described below.

According to Kerner (2002), wide moving jams can, however, emanate within a queue of synchronized flow (at locations that are well upstream of bottlenecks). Discharge flow from these moving jams can be lower than the flow at the downstream region of synchronized flow. With this flow difference, traffic becomes freely flowing at the downstream end of the jam; and the free flow state propagates forward toward the bottleneck, eventually dissipating the queue there. This pinch effect is illustrated in Figure 2.



**Figure 2 Illustration of pinch effect**

Phenomena very similar to this pinch effect have been observed at the study site (See section 3.3). The objective of this current work has been to meter on-ramps to affect these phenomena to favorable ends. A review of previous freeway metering experiments is therefore furnished next.

## **2.2 On-ramp metering experiments of freeway bottlenecks**

Edie and Foote (1960) carried out a field study to increase traffic discharge flow through a tunnel based on kinematic wave theory. Vehicle platoons were restricted in size by introducing gaps between vehicles at the upstream end of the tunnel. In this way, overloading the tunnel with traffic was avoided. This control method was reported to increase the tunnel's discharge flow by preventing queues from forming inside the tunnel.

Building on this earlier work, Wattleworth and Berry (1965) postulated that on-ramp metering can prevent queues from forming at a freeway bottleneck and thus maximize output flow from the freeway by storing a freeway bottlenecks' excess demand on the upstream on-ramps. They also noted that the goal of ramp metering should be to maximize output flow from the freeway because all else equal, this goal is equivalent to minimizing total travel time spent in the freeway. On the other hand, Atol and Bullen (1973) argued that ramp metering should extend uncongested time periods by limiting a bottleneck's discharge flows to the extent that the probability of capacity drop is maintained below a certain threshold. Such a threshold approach inevitably requires more or less "over-control" to prevent capacity drop at a bottleneck.

Other researchers (Newman et al, 1969; Allen and Newell, 1976) noted that restricting on-ramp flows can create long ramp queues and thereby induce some drivers to divert from the freeway to alternative routes along local streets, for example. Cassidy (2003) pointed out that such diversion can either increase or decrease commuter delay in the whole network of freeway and local streets; and that the evaluation of the diversion effect is difficult to carry out in the field. For this reason, diversion was not explored in the present research. The study site, moreover, enjoys few direct alternative routes.

Unfortunately, we have few ramp metering field studies that conclusively report success in increasing outflows from the freeway with serial bottlenecks despite methodological advances of on-ramp metering (Consult with the comprehensive reviews of Lovell, 1997 for many innovative analytical techniques). Review of the relevant previous field studies is provided below.

Some studies (Papageorgiou et al, 1998; Diakaki et al, 2000) reported that flow and speed within freeway serial bottlenecks increased if on-ramp flows are metered. However, these studies, as noted by Cassidy (2003), do not necessarily indicate that on-ramp metering diminished system-wide commuter delays. This is because higher flow on internal links with a freeway system might be realized due to the transfer of commuter delays from the freeway to its on-ramps and connecting surface streets.

Other corridor-wide field studies (Haj-salem et al, 1995, MnDOT, 2001, and Ahn et al, 2007) used savings in drivers' overall trip time as a performance measure of metering strategies that were deployed at multiple on-ramps along corridors. However, these

studies did not focus on the capacity of serial freeway bottlenecks; and did not confirm that the bottlenecks were unaffected by queues from downstream.

Cassidy (2003) argued that on-ramp metering can be detrimental if it's not tailored to remedy the cause of the bottleneck. As an example, he considered a situation in which freeway queues are caused by vehicles competitively exiting toward a freeway off-ramp. He pointed out that an on-ramp metering scheme, if it's myopically applied, might cause more system wide delay at this kind of diverge bottleneck by penalizing the vehicle which is not destined to the problematic off-ramp while favoring the vehicles bound for the off-ramp. In light of this, he suggested that designing on-ramp metering schemes should be preceded by pinpointing the cause of a bottleneck using high-resolution analytic tools such as curves of cumulative vehicle counts.

Finally, Cassidy and Rudjanakanoknad (2005) field-tested a metering strategy to reverse the capacity drop phenomenon at an isolated active on-ramp merge bottleneck. (Notably, the study verified that the bottleneck was indeed an active one.) Field experiments revealed that capacity drops were recovered if restrictive and relaxed on-ramp metering rates were alternatively deployed in response to the level of queuing upstream of the bottleneck. The researchers also reported that metering an on-ramp in this manner can increase long-run average discharge flow from the bottleneck during a rush, although each recovery in the bottleneck's discharge flow was sustained for only a limited time (i.e. always less than 13 minutes). It remains to be seen if metering strategies can be made more robust and be used to benefit freeways with serial bottlenecks.

### **3. Observed traffic phenomenon at the site**

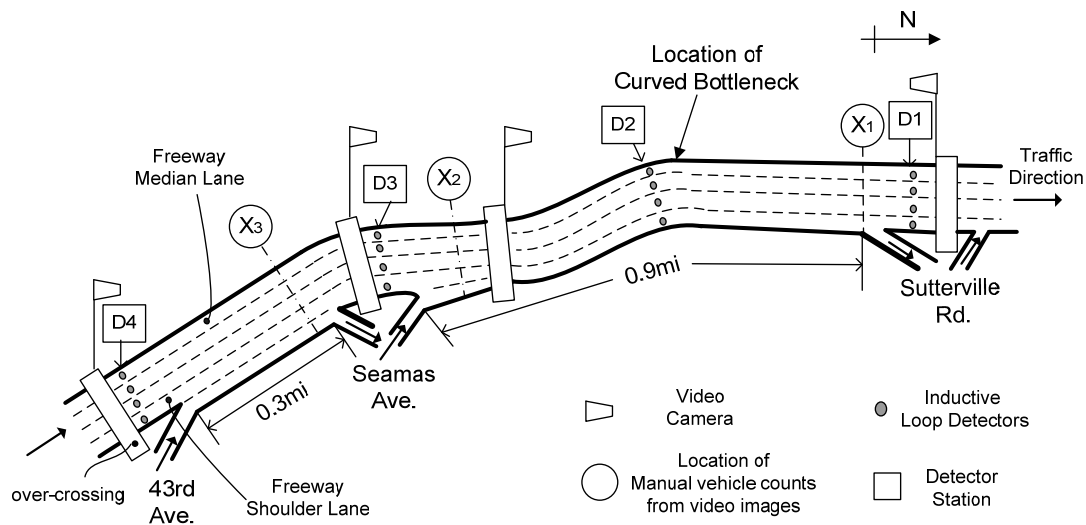
This section describes the mechanisms observed in which discharge flows at the study site are affected by the queues at the site's serial bottlenecks. The site contains a downstream curved section that frequently becomes an active bottleneck during much of the morning rush (Sec 3.1). The discharge flows from this curved section are characterized by sequences of reductions followed by recoveries; the flow reductions are short-lived early in the rush but later in the rush become longer-lived. This creates significant reductions in the bottleneck's long-run average discharge flow and thus



increases system-wide commuter delays (Sec 3.2). These damaging long-lived reductions in outflow late in the rush are found to be caused by the pinch effect which occurs whenever a queue from the curved bottleneck grows long and spills-over to an upstream weaving section near the interchange of 43rd Ave. The spill-over queue severely constrains flows in this upstream weaving section; so much so, that the spill-over queue gradually dissipates due to lack of discharge flows from the weaving section, i.e. the queue's tail propagates from the weaving section forward toward the curved bottleneck downstream. As the queue's tail approaches the curved bottleneck, its discharge flow recovers (Sec 3.3). Whether or not a queue from the curved bottleneck becomes damaging by growing long depends to some extent on the inflows at the on-ramps upstream of the curved bottleneck (Sec 3.4). Finally, lessons learned from the above observations are summarized in their relevance to ramp metering (Sec 3.5).

### 3.1. Freeway site and its curved bottleneck

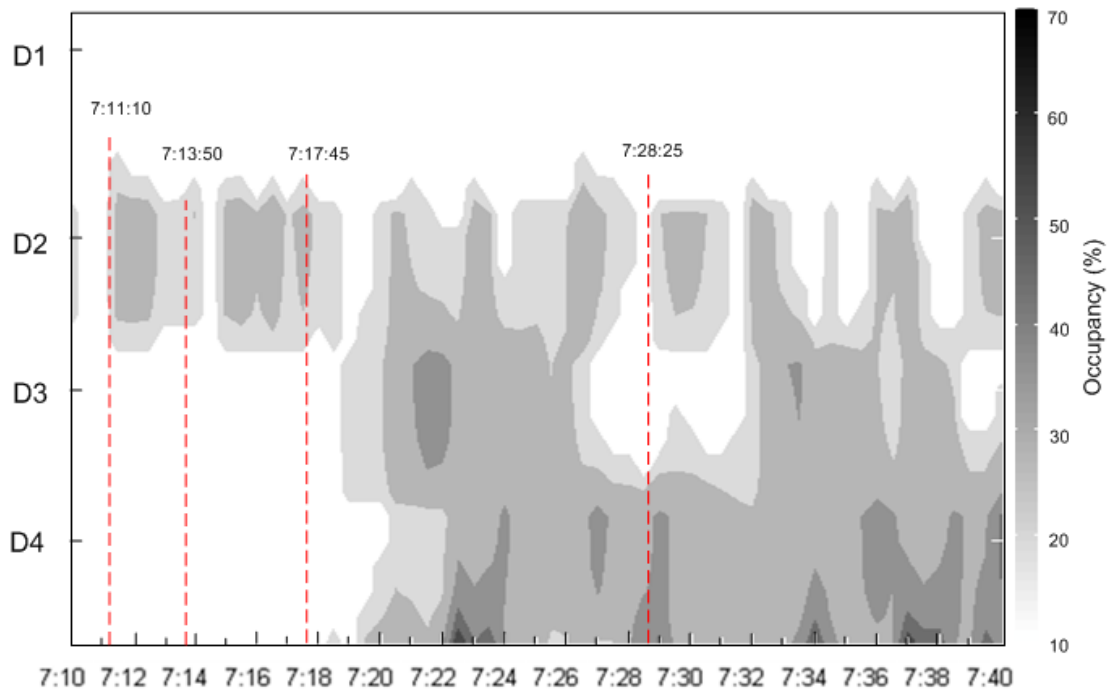
Figure 3 illustrates the study site, a stretch of northbound Interstate 5 in Sacramento, California. During much of morning rush periods, a bottleneck activates in the downstream curved section (near detector D2). Bottlenecks also form upstream, at the merges created by the on-ramps for Seamas Ave and for 43<sup>rd</sup> Ave.



**Figure 3 Study Site, Northbound I-5, Sacramento, California**

The data used in this study were collected both from loop detectors and from video cameras. The former provided aggregated vehicle counts and average occupancies per 30 sec sampling intervals for all freeway lanes at D<sub>1</sub>~D<sub>4</sub> (see Figure 3). Because no detector data are available at on-ramps and off-ramps, vehicle counts on the ramps were manually extracted from videos. In addition, vehicles were counted from videos in all freeway lanes at the locations labeled X<sub>1</sub>~X<sub>3</sub> in Figure 3.

The presence of the curved bottleneck is confirmed for one morning rush (on November 18, 2006) using the occupancy contour map in Figure 4. Note that detector occupancies of about 17% or less (non-shaded regions) denote free flow traffic conditions (flow=demand); and that occupancies greater than about 17% (shaded regions) denote queues. Figure 4 thus reveals that the freeway location near detector station D<sub>2</sub> exhibits free flow conditions downstream with frequent queuing upstream; i.e. the curved section periodically becomes an active bottleneck. Note too from the figure that queues almost always persist somewhere upstream of the curved section from 7:15 onward.



**Figure 4 Occupancy contour map (10/18/06)**

### 3.2 Time-varying discharge flows from the curved bottleneck

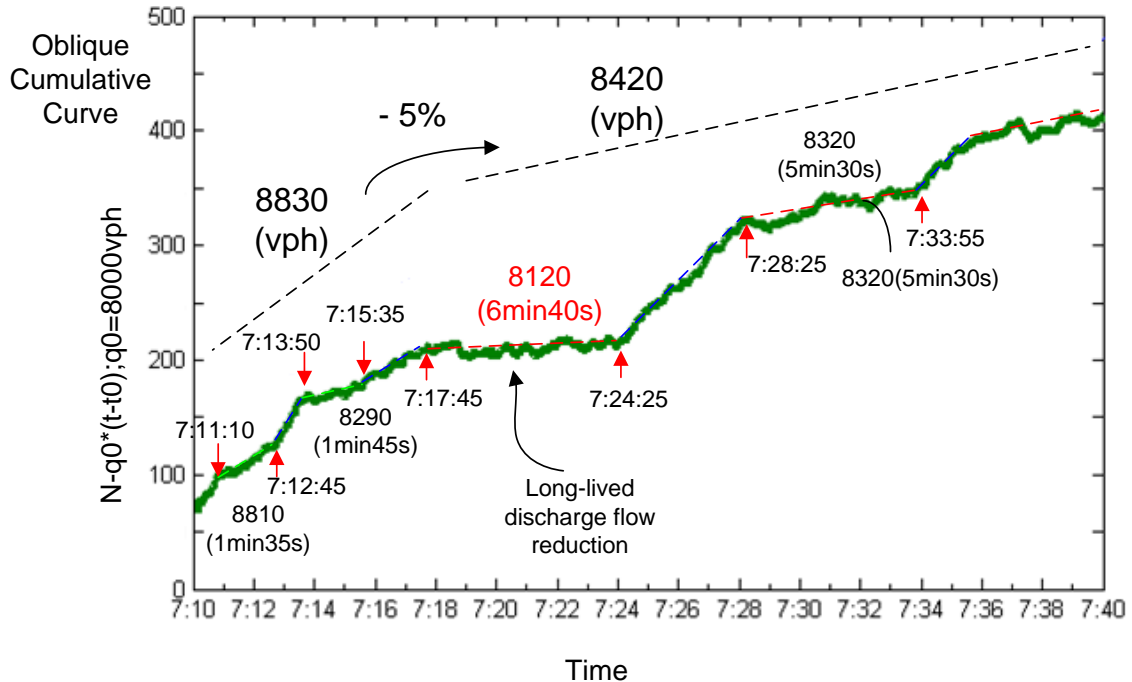
We next examine the discharge flows from the curved section using Figure 5. It presents a curve of cumulative counts of vehicles passing downstream location X1 in oblique coordinates (O-curves) that were extracted from videos. Note that the slopes of an O-curve present flows in excess of a background flow reduction, which was 8000 vph in the present case.

Figure 5 reveals that the flows departing from the freeway's curved section are characterized by sequences of reductions followed by recoveries<sup>3</sup>. Note from the dashed lines drawn through the O-curve that short-lived flow reductions began at 7:11:10 and at 7:13:50. Note too the longer lived reductions beginning later in the rush at 7:17:45 and at 7:28:25. These later events result in reductions in long-run average discharge flow of 5%, as highlighted by the dashed lines drawn above the O-curve in the figure. These reductions in long-run average discharge flow create additional commuter delays at the freeway site (See Cassidy, 2003 for more detailed discussion on this issue).

Tellingly, the short-lived flow reductions that occurred early in the rush (Figure 5) roughly coincide with periods marked by short queues just upstream of the curved bottleneck (Figure 4). Moreover, the longer-lived flow reductions that occurred later in the rush, roughly coincide with long queues that persistently swamped the upstream portion of the freeway site.

---

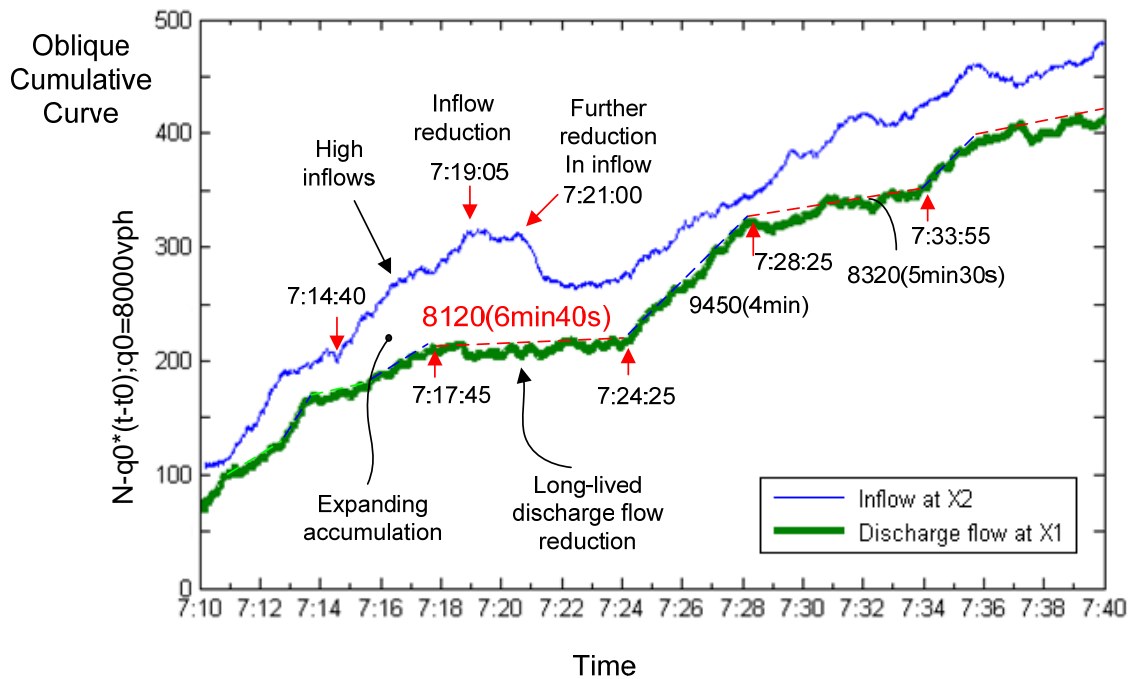
<sup>3</sup> These flow reductions were no doubt triggered by random events, such as a lane-change maneuver performed by a truck with low acceleration capabilities; see Laval and Daganzo (2006)



**Figure 5 Discharge flow from the curved bottleneck (10/18/06)**

### 3.3 Bottleneck interactions

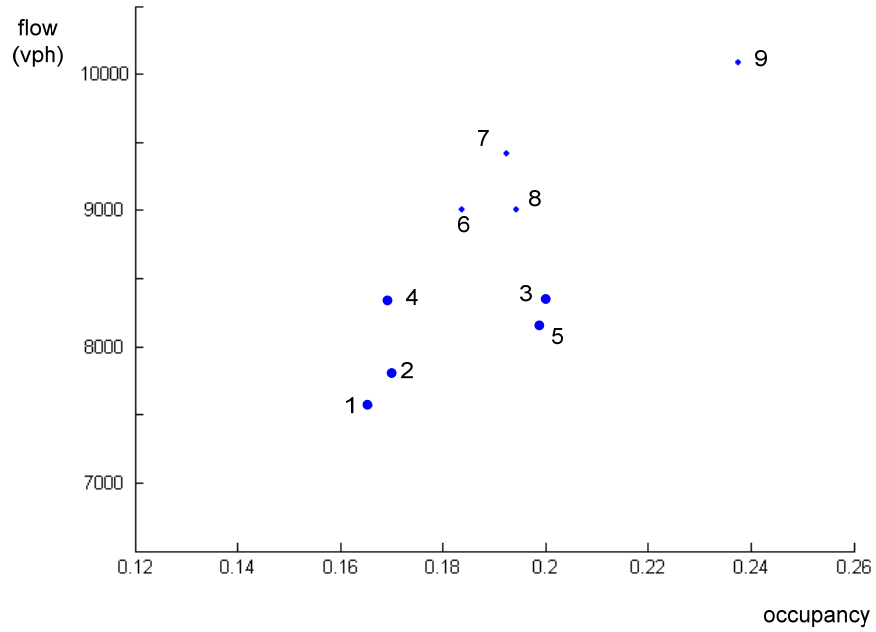
Here we show how the long queue from the curved bottleneck interacts with merge (and weave) bottlenecks upstream to produce long-lived reductions in discharge flow. We examine first the O-curves in Figure 6; these present the discharge flows from the curved section measured at X1 (boldface curve in the figure) and its input flows measured upstream at location X2 (light curve). Note the long-lived discharge flow reduction that began at 7:17:45. The steep slope of the thin curve in Figure 6 reveals that high inflows persisted during this time; and that as a result, the queue from the curved bottleneck grew long (as evident from the expanding accumulation between the curves). The figure further reveals that this long queue eventually spilled-over to the merge at Seamas Ave. and constrained the flow there: note the sudden input flow reduction at 7:19:05.



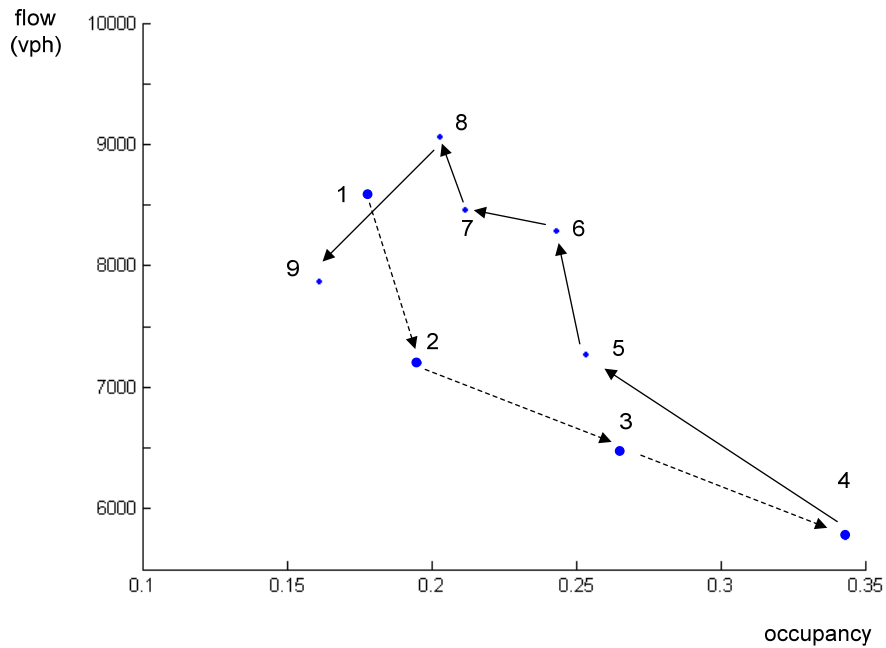
**Figure 6 Inflow and discharge flow of the curved section (10/18/06)**

This spill-over set into motion the pinch effect that eventually caused the discharge flows from the curved bottleneck to recover. The bad news, however, is that the mechanism of this pinch effect was slow-moving; such that the discharge flow did not recover until 7:24:25; i.e the flow reduction was long-lived. A clue concerning the nature of this slow-moving mechanism is evident in the further reduction in input flow at 7:21 shown by the thin curve in Figure 6. Greater insights into the mechanism come by examining flow and occupancy measured by loop detectors as we will do next.

Figure 7 and 8 present time series of flow and occupancy that were measured at detector D2 and D3, respectively. Each data point, which is numbered in chronological order, is the average flow and occupancy per 1 minute time interval.



**Figure 7 Flow-occupancy plots at detector D2 (7:18~7:26; 10/18/06)**

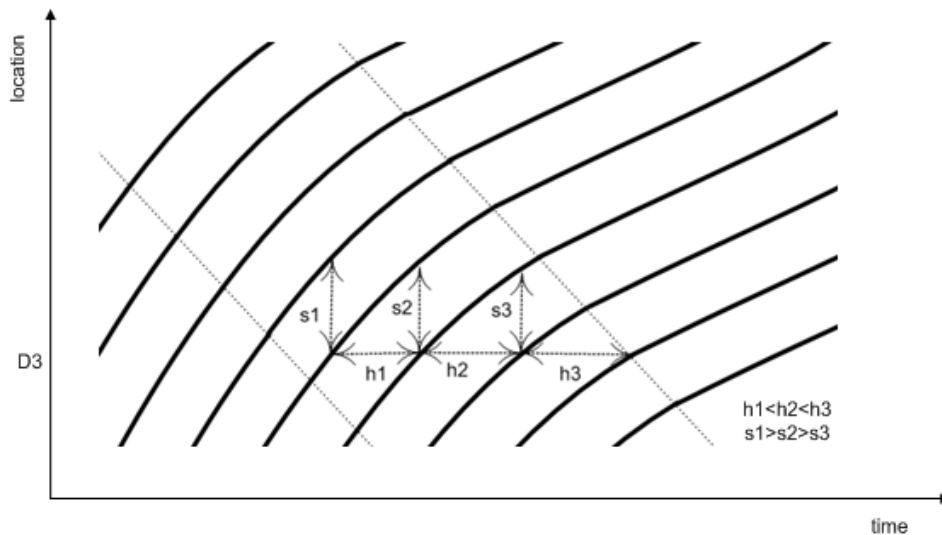


**Figure 8 Flow-occupancy plots at detector D3 (7:18~7:26; 10/18/06)**

We begin with Figure 7. The early half of the data points (1~5) shown with bold circles at D2 show that there was low discharge flow from the curved section. This low

output flow was due to queue upstream. This queue became long and eventually arrived at D3. The long queue's arrival is revealed by gradual flow reductions coincident with gradual increases in occupancy, as shown by the early data points (1~4) in Figure 8. Interestingly, change in flow and density was gradual at D3 (i.e. speed went down slowly over time). Such a gradual speed reduction implies that there was a transition zone at the back of the queue. The transition zone can be illustrated by set of (fictitious) vehicle trajectories near detector D3 as described below.

Trajectories shown in Figure 9 present the way that drivers gradually reduce their speeds when they run into the back of a queue. As a result, the spacing of vehicles (which are proportional to the inverse of occupancy) gradually diminished while the headway of vehicles (the inverse of flow) gradually increased at the detector location D3.



**Figure 9 Illustration of transition zone at the back of queue**

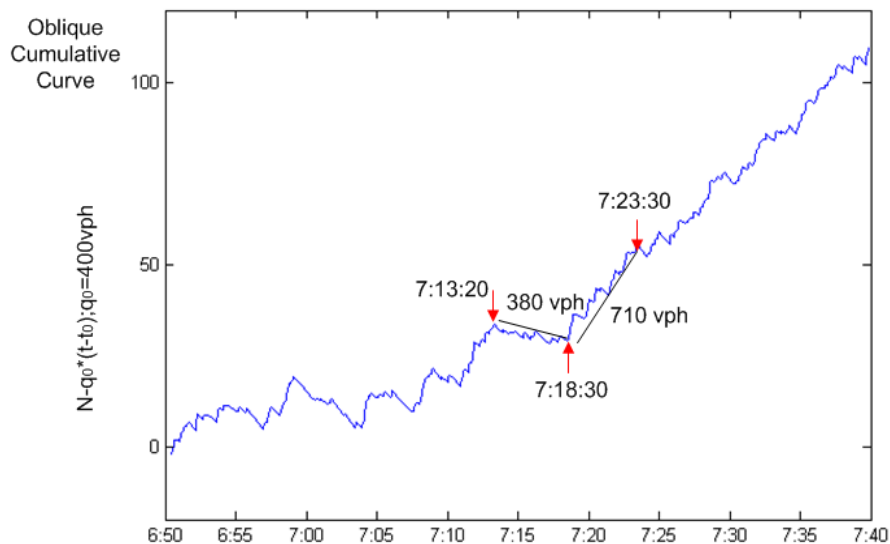
The later data points (5~9) at the detector D3 reveal the mechanism by which the freeway's long queue was dissipated (See again Figure 8). After a new bottleneck became pinched at the weaving section near the 43rd Ave interchange, low discharge flow from the pinched bottleneck starved the downstream queue. As a result, the flow gradually increased at detector D3 while occupancy simultaneously and gradually diminished there. This implies that speed went up gradually, which can be explained by the same logic that was used for gradual speed reductions (Figure 9).

Interestingly, as the queue shrinks downstream of the 43<sup>rd</sup> Ave interchange, drivers in the queue adopt short headways. This behavioral change resulted in the rise of occupancy coincident with the increase of flow at detector D2 (See Figure 7 and refer to the data points 5 ~ 9).

All the mechanisms illustrated above were recurrently observed within and across days. It turns out that on-ramp inflows from Seamas and 43<sup>rd</sup> Aves were influential in initiating the long-lived reduction of the curved bottleneck's discharge flow. This influence will be verified by detailed analyses in the next section.

### 3.4 The effect of on-ramp flows on the curved bottleneck's queue

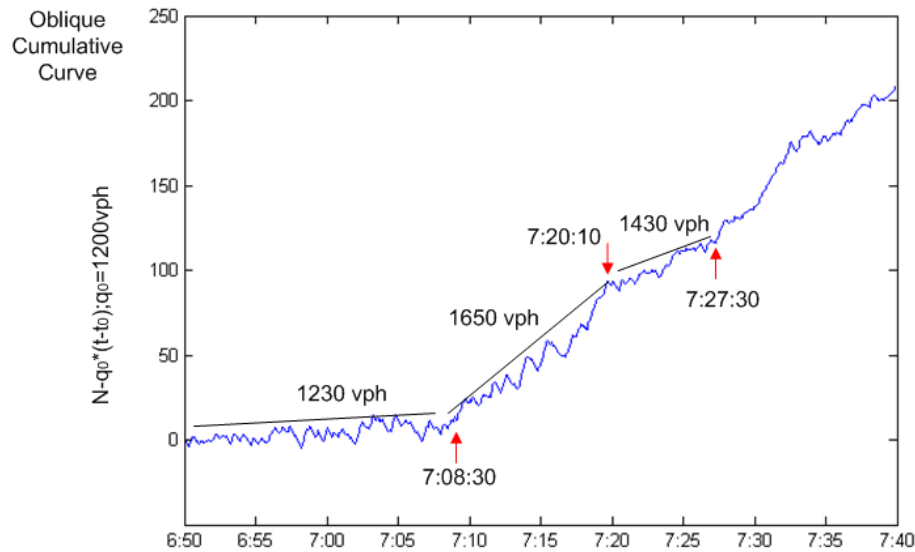
We show below that on-ramp inflows at Seamas Ave play a key role in the growth of the curved bottleneck's queue. These on-ramp inflows at Seamas suddenly increased (from 380vph to 710vph) at 7:18:30 (See Figure 10). This caused the bottleneck's queue to propagate beyond X2 and restrict the flows there (see again the thin curve in Figure 6 and note that the flow restriction at X2 began at 7:19:05). Data further show that the high on-ramp flows at Seamas made the queue in the freeway shoulder lanes especially dense at locations upstream of X2.



**Figure 10 On-ramp flows from Seamas Ave. (10/18/06)**



We can also infer that high on-ramp flows at 43<sup>rd</sup> Ave further elongated the curved bottleneck's queue. On-ramp flows at 43<sup>rd</sup> Ave. suddenly increased (from 1230vph to 1650vph) at 7:08:30 as seen in Figure 11. High on-ramp flows at 43<sup>rd</sup> Ave prevailed until the curved bottleneck's long queue arrived at detector location D4 and thus constrained on-ramp flows at 43<sup>rd</sup> Ave. (See Figure 11 and note that on-ramp inflows at 43<sup>rd</sup> Ave were reduced from 1650 vph to 1430 vph at 7:20:10, which is when the freeway queue arrived to D4.)



**Figure 11 On-ramp flows from 43rd Ave. (10/18/06)**

### 3.5 Lessons learned from empirical study

The lessons learned from the above study can inform our choice of a ramp metering strategy. When a queue first appears at the curved bottleneck, we can attempt to dissipate the bottleneck's queue as soon as possible with somewhat restrictive metering. If the freeway queue disappears as a result, we can then relax the metering to avoid over-control and to flush the on-ramp queues. In addition, once a queue becomes non-transient and starts growing long, we can try harder to dissipate the queue via more restrictive metering. This more restrictive metering will at times fail due to persistently high traffic demand from the freeway mainline. As a result, the pinch effect will eventually occur

when freeway queue reaches the weaving section near the interchange of 43<sup>rd</sup> Ave. We will then let the pinch effect take its course by maintaining restrictive metering at both on-ramps to avoid further queue growth downstream of the 43<sup>rd</sup> interchange. After the pinch effect clears the curved bottleneck's queue, the above procedures can be repeated, as needed.

## **4. Ramp metering field tests**

This section describes a proposed ramp metering scheme and the outcome from testing it at the study site. The lessons learned from the empirical study were translated into a set of metering rules to combat the queues from the curved bottleneck (Section 4.1). To meter on-ramp inflows in response to changing traffic conditions, the curved bottleneck's queues must be detected and tracked via data from study site's (existing) loop detectors. The proposed indicators are time series of occupancies (Sec 4.2). These indicators dictated the metering rates at the on-ramps for 43<sup>rd</sup> and Seamas Avenues during the field tests; and due consideration was given to on-ramp queue lengths (Sec 4.3). The field tests were successful: the curved bottleneck's discharge flows increased to the extent that queues diminished upstream of the bottleneck (Sec 4.4).

### **4.1 Proposed ramp metering scheme**

The scheme aims to reduce (preferably dissipate) the curved bottleneck's queues as much as possible by restricting on-ramp inflows at Seamas and 43<sup>rd</sup> Aves. Metering rates deployed at both on-ramps vary depending on the queue lengths that are detected both on the freeway and on the ramps.

In response to the short queues that formed at the curved bottleneck during early portions of the morning rush, a very restrictive metering rate (300 ~ 400vph) was deployed at the Seamas on-ramp, with a less restrictive metering rate (1500 ~ 1600vph) at the 43rd on-ramp. We deploy relaxed metering rates at both the Seamas and 43rd on-

ramps (700 and 1600vph, respectively) if the freeway queue disappears as a result of the restrictive metering at Seamas.

On the other hand, queues from the curved bottleneck tend to grow long during much of the later rush due to persistently high traffic demand from freeway's mainline. Thus whenever possible, both the Seamas and 43rd on-ramps are metered at rates of 300~400vph and 1100~1300vph, respectively. However, the metering rates are relaxed whenever the on-ramp queues grow excessively long.

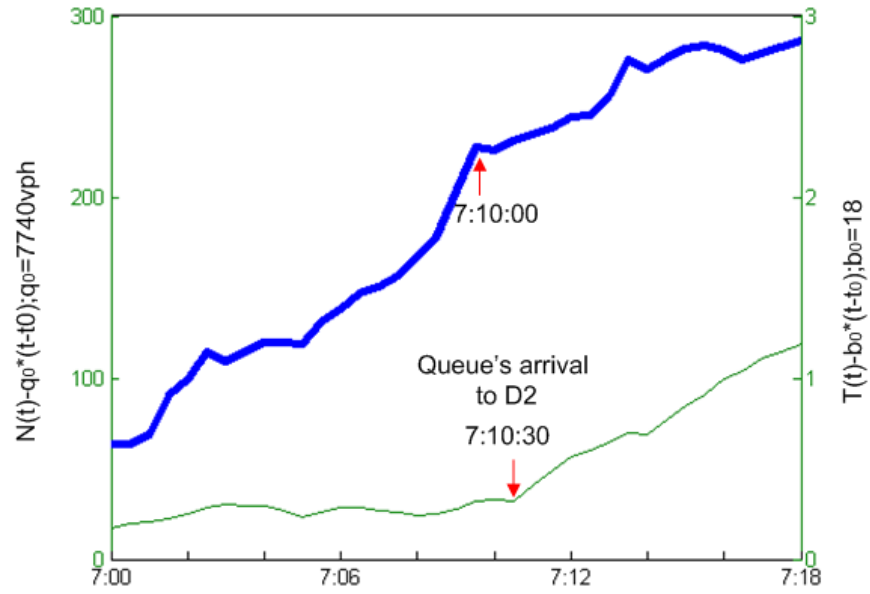
On some occasions, the tail of the freeway queue remains near the interchange of 43<sup>rd</sup> Ave after the pinch effect completely clears the downstream queue. In these cases, a very relaxed metering rate (700vph) was deployed at the Seamas on-ramp, while restrictive metering rate was maintained at the 43<sup>rd</sup> on-ramp. The relaxed metering at Seamas was maintained until a queue formed again at the curved bottleneck.

## **4.2 Indicators to trace the curved bottleneck's queue**

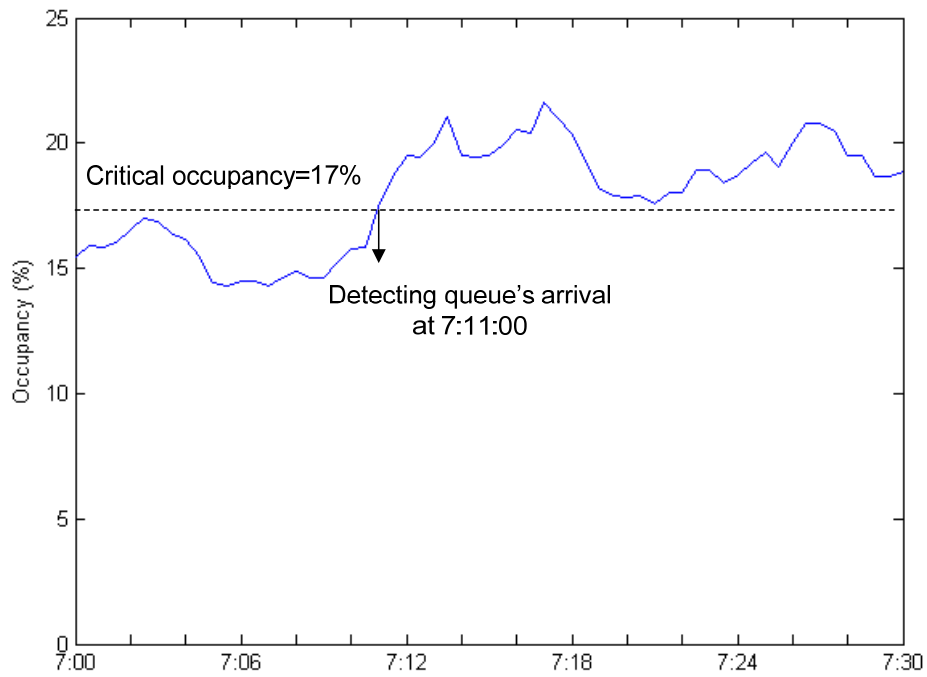
Indicators are required to detect when: (1) a queue forms at the curved bottleneck, and (2) that queue grows long and reaches the upstream merges at Seamas and 43rd. Time series of occupancies across all lanes were used for this purpose. A 3-minute moving average was used to filter out fluctuations. The effectiveness of the indicators is examined below.

The indicators identify the queue's arrival to the detectors D2 and D3 within a minute of the actual occurrences. Figure 12 presents an O-curve (boldfaced) and a cumulative occupancy curve, i.e. T-curve (thin) measured at detector D2. These curves reveal that the queue arrived to the detector D2 at precisely 7:10:30: note how the slope of O-curve declines simultaneous to an increase in the slope of T-curve, and this is the signature of a queue's arrival to a detector (Cassidy and Bertini, 1999). Figure 13 shows that the occupancy at detector D2 exceeds the critical value of 17% at 7:11:00; i.e. the time series identified the queue's arrival to detector D2 to within 30 seconds in this case.

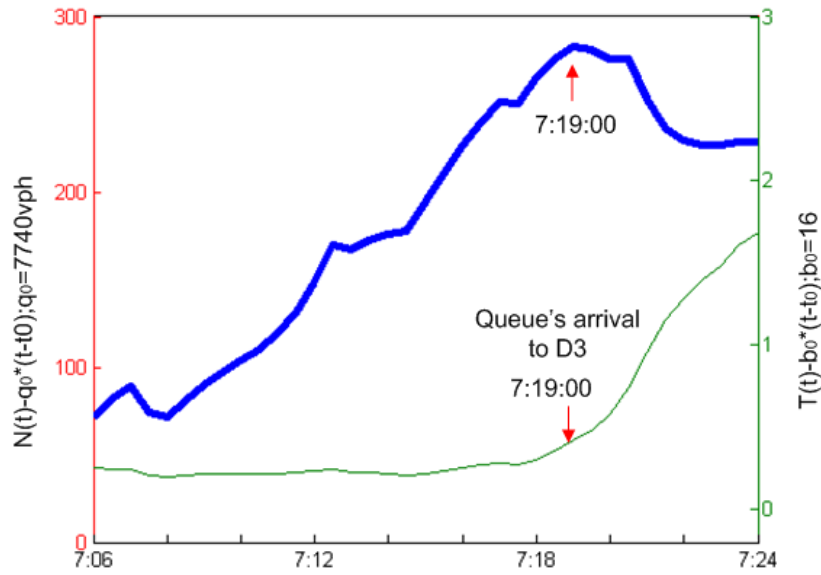
Similarly, the indicators determine approximately when the long queue reaches detector D3. An O-curve and T-curve in Figure 14 show that the curved bottleneck's queue arrived to D3 at 7:19:00; while the time series of occupancies in Figure 15 detect this event to within 30 seconds.



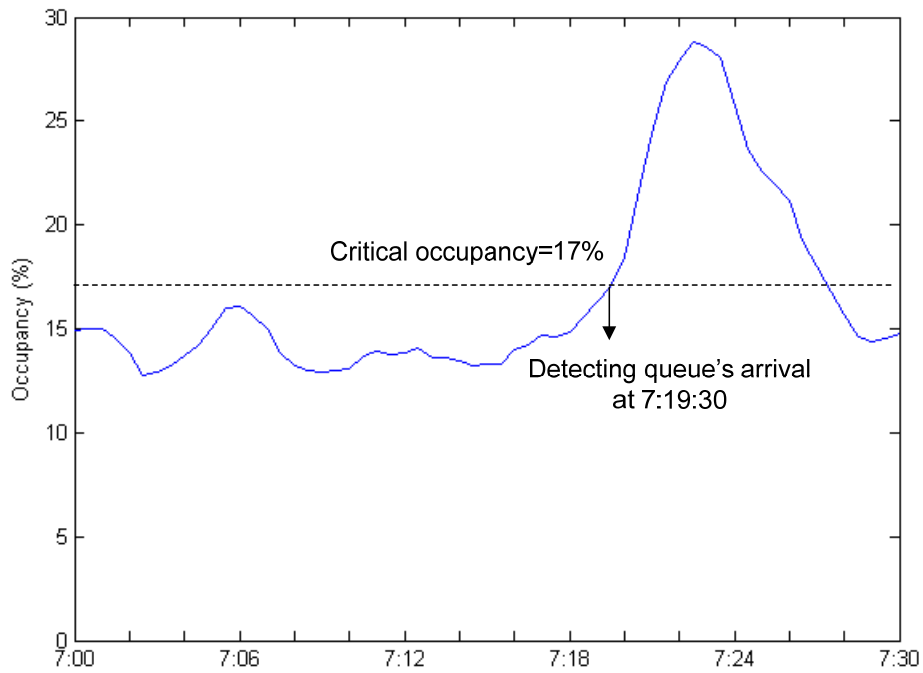
**Figure 12 O-curve and T-curve at D2 (10/18/06)**



**Figure 13 Time series of occupancies at D2 (10/18/06)**  
**(3-minute moving average)**



**Figure 14 O-curve and T-curve at D3 (10/18/06)**



**Figure 15 Time series of occupancies at D3 (10/18/06)**  
**(3-minute moving average)**

### **4.3 Field experiments**

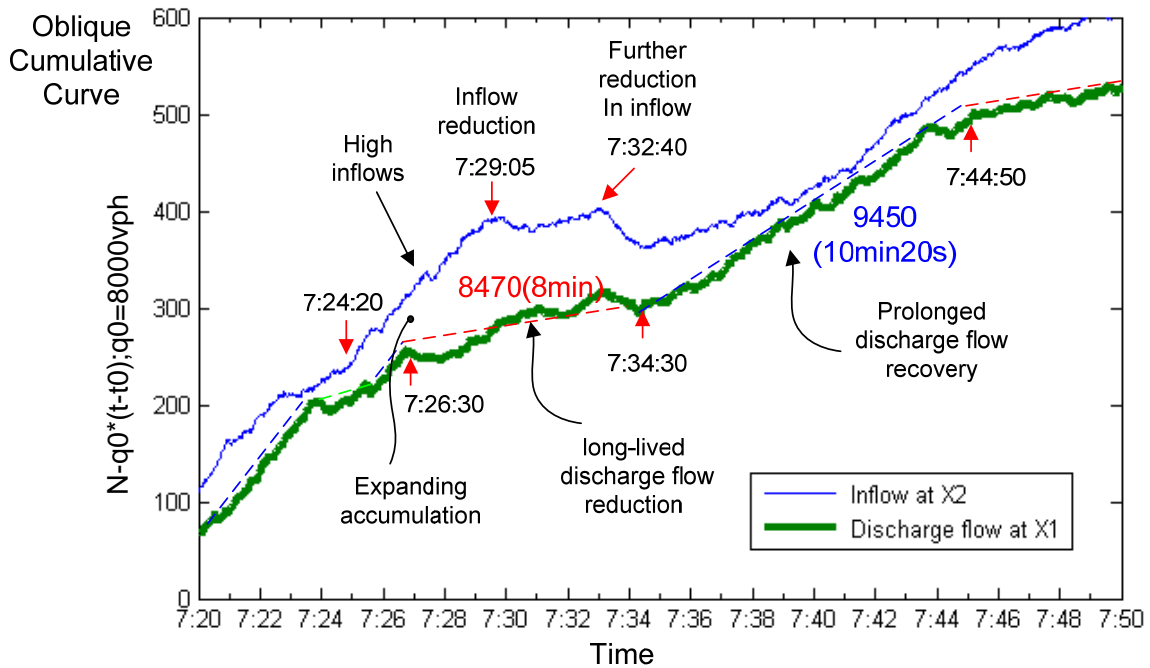
Field experiments were conducted during morning rush periods on 9 weekdays in spring and fall 2009. The proposed metering scheme was implemented remotely at the traffic control center. A human operator at this center instituted on-ramp metering rates at Seamas and 43<sup>rd</sup> Aves based on the queue lengths that were detected, both on the freeway and on the ramps.

As regards the freeway queue, metering rates were relaxed both at the Seamas and 43<sup>rd</sup> Aves when occupancies at detectors D2 and D3 were below 17% (i.e. when traffic between D2 and D3 was freely flowing). As the freeway queue from the curved bottleneck arrived to detector D2 and thus increased occupancies above 17% at that location, on-ramp inflows were restricted at Seamas Ave. In some instances, this restrictive metering was not enough to contain the growth of the freeway queue. Thus, occupancies at the upstream detector D3 eventually surged above 17% when the queue arrived to that detector. In this case, very restrictive metering rates were deployed both at the Seamas and 43<sup>rd</sup> Aves.

In addition, human observers were deployed to track the length of queues at the two on-ramps. When either on-ramp became filled with queued vehicles, the ramp's observer called the operator in the traffic center to request that the metering rate be relaxed. The operator decided whether to relax the rate depending on the length of freeway queue. The effects of the scheme on the study site's traffic are discussed next.

### **4.4 Results**

We first present the traffic details that were observed during the conduct of the field experiments. O-curves in Figure 16 show inflow and discharge flow that were measured at X2 and X1, respectively, during a typical morning rush (on May 14, 2009) when the proposed metering scheme was deployed. For comparison, see also the O-curves in Figure 6 that show traffic details that were observed in the absence of metering (in 2006).



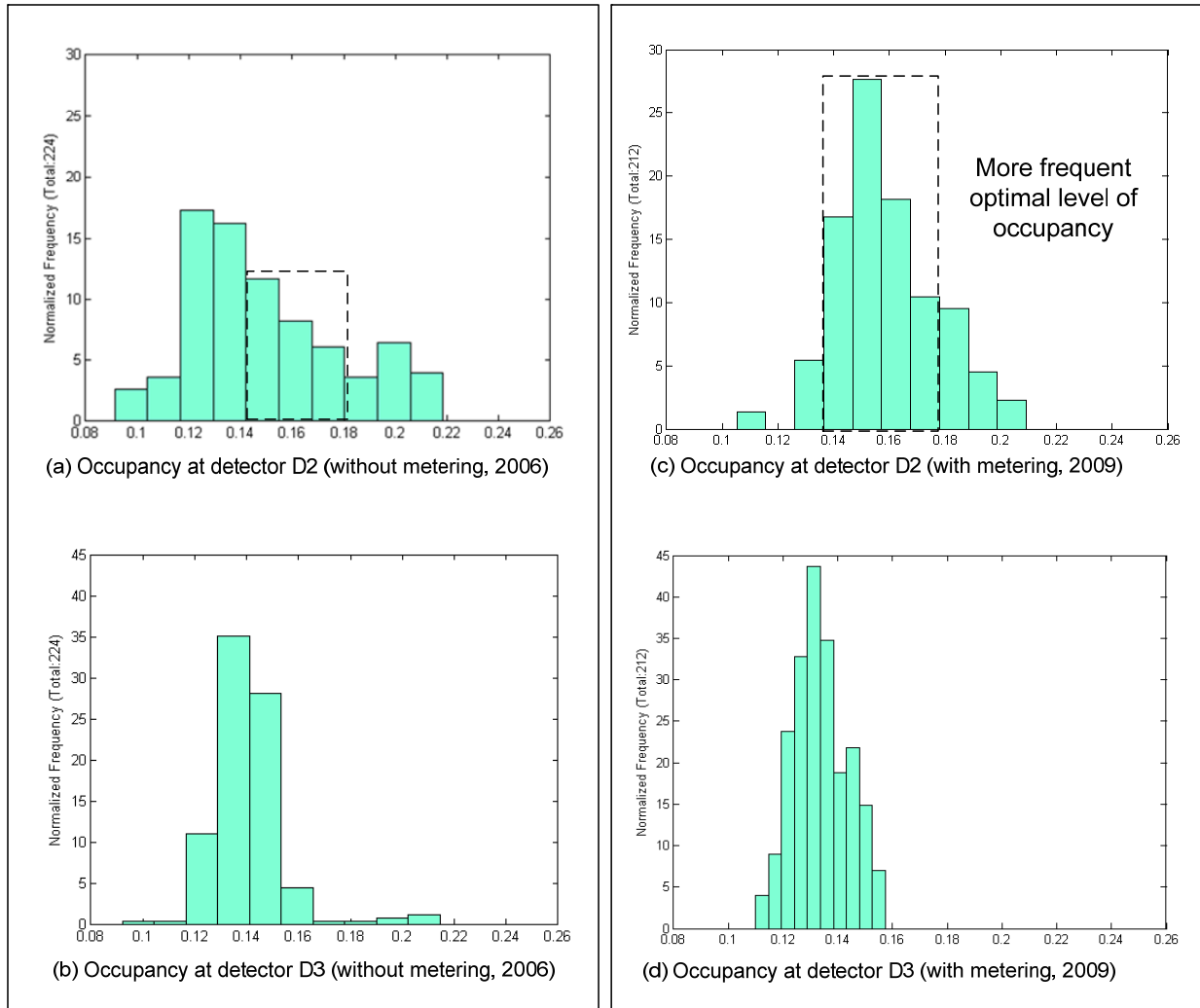
**Figure 16 Inflow and discharge flow from the curved section (05/14/09)**

During the field experiments, long-lived reductions in discharge flow were less severe in magnitude. This is confirmed by the reduced discharge flow (8470 vph) in Figure 16 which is substantially higher than its counterpart (8120 vph) previously shown in Figure 6. It seems that restrictive metering at the Seamas on-ramp reduced the frequency of disruptive lane changing maneuvers that occurred just upstream of the curved section (Consult Cassidy et al, 2009 for discussion on this so-called smoothing effect that arises from reductions in disruptive lane changing).

Moreover, discharge flow recovery in Figure 16 (10min 20s) was sustained for a longer period than for its counterpart in Figure 6 (4min). This prolonged recovery in discharge flow was enabled by restrictive on-ramp metering at both the Seamas and 43<sup>rd</sup> on-ramps. (The restriction here was performed to dissipate the queue that forms again at the entrance of the curve after the pinch effect cleared the queue upstream of the curve.)

The above favorable effects of on-ramp metering were reproducible within the day and across experiment days. Furthermore, analyses of traffic data aggregated across experiment days also confirm the benefit of on-ramp metering, as presented below.

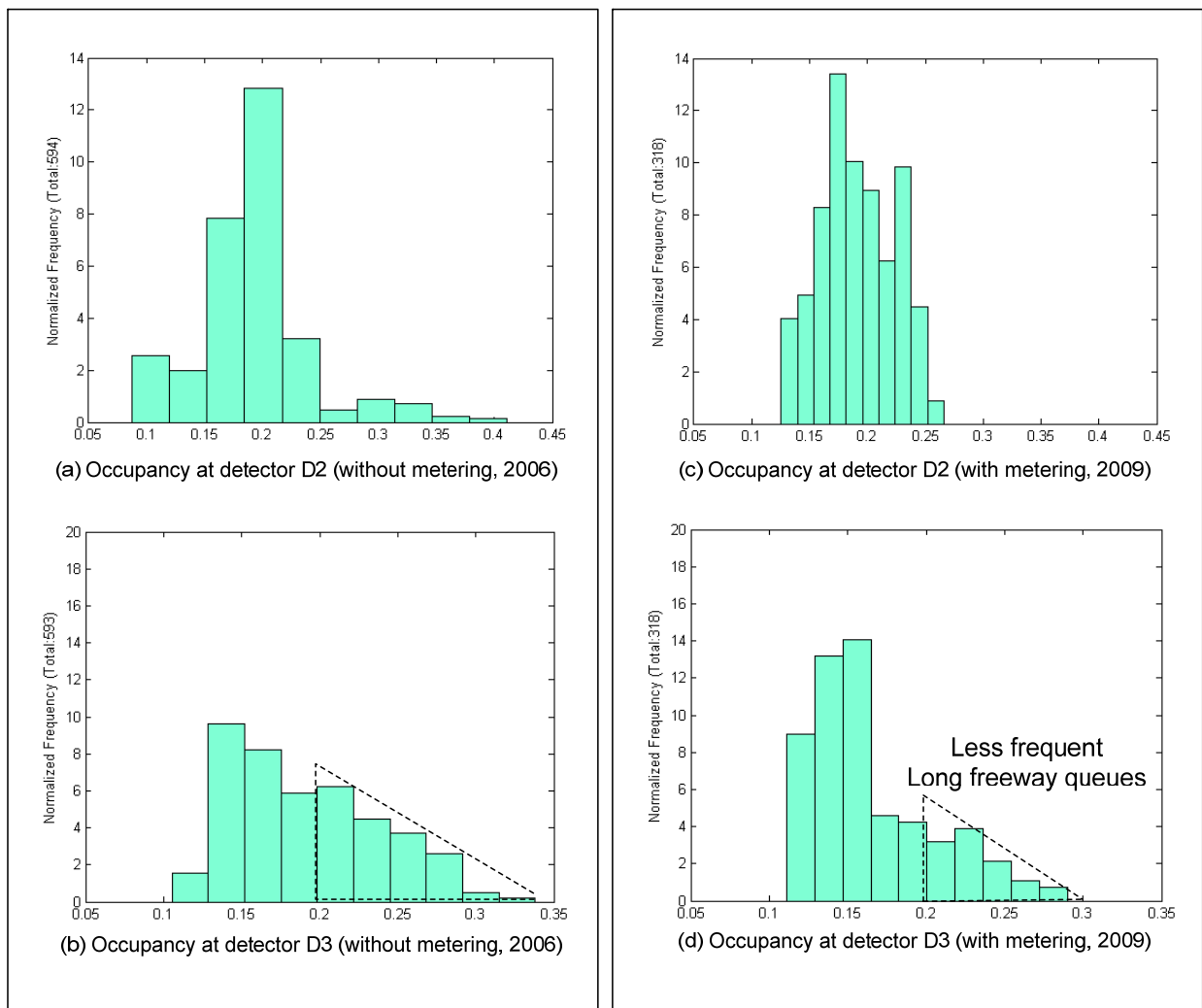
Occupancies at detector D2 and D3 are presented as proxies of densities upstream of the curved bottleneck with and without on-ramp metering (Refer to the histograms of the occupancies in figure 17 and 18 for early and late rush, respectively). Recall that there was no metering in 2006; and that our proposed metering scheme was deployed at the Seamas and 43<sup>rd</sup> on-ramps in 2009. Data from five weekdays were used to construct the histograms for each year.



**Figure 17 Histograms of occupancies measured at detector D2 and D3 (Early rush)**



As shown in (b) and (d) of Figure 17, low occupancies (i.e. occupancies below 17%) were observed at upstream detector D3 during the early rush, both in 2006 and 2009. These low occupancies at D3 confirm that the curve bottleneck's queue rarely grew long in the early rush. More importantly, histograms (a) and (c) in Figure 17 reveal that occupancies measured at detector D2 in 2009 were more frequently distributed near 17% (optimal occupancy in terms of curve bottleneck's discharge flow) than those measured in 2006.



**Figure 18 Histograms of occupancies measured at detector D2 and D3 (Later rush)**

Histograms also show that occupancies were higher than the critical occupancy 17% at upstream detector D3 during much of the later rush, both in 2006 and 2009 (See Figure 18 (b) and (d)). These higher occupancies at D3 were largely caused by the long freeway queues from the curved bottleneck. Notably, we observed high occupancies at D3 less frequently in 2009 than in 2006. This occupancy change at D3 indicates that queues diminished upstream of the curved bottleneck during the morning rush periods when on-ramps were metered.

All the above-cited changes favorably affected the long-run average of the bottleneck's discharge flow as shown in Table 1. In the early rush, we observed 7 % higher discharge flow from the curved bottleneck when the two on-ramps were metered as per our scheme. (The measured discharge flows were sustained for 20 minutes on average.) In addition, we observed a 5% increase in the average discharge flow during the later rush. (The rates were sustained for 60 minutes on average) On the whole, 5.5% increase occurred in average discharge flow. Coarse analysis indicates that for present-day demand, an outflow increase by 5.5% would diminish morning delay by roughly 30%. (This would translate to a reduction of about 430 vehicle-hours each morning rush.)

**Table 1. Long-run average of discharge flow from the curved bottleneck**

Ramp Metering	Year	Average <sup>1)</sup> discharge flow (vph) during morning rush periods <sup>2)</sup>					
		Early rush (20min)	$\Delta$ (%)	Later rush (60min)	$\Delta$ (%)	Total (80min)	$\Delta$ (%)
None	2006	8760	7	8320	5	8430	5.5
Proposed	2009	9410		8750		8920	

1) The average of five weekday morning rush hours for each year

2) Morning rush periods when exogenous queues did not spill -back from downstream

Furthermore, our study shows that our proposed metering scheme did a slightly better job of increasing the curved bottleneck's average discharge flow during the morning rush than did the metering that Caltrans has deployed at the study site (The difference in this flow was 2.3%). This result points to the benefits of metering on-ramps in response to the length of a bottleneck's queue.

The proposed metering scheme may not be due full credit for increasing the discharge flow. There has been an economic downturn since early 2008, which has generated lower traffic demand at the site. With the aid of the exogenous effect, the curved bottleneck's queues could be more easily controlled. The point, however, is that whatever the means may be, 1) higher long-run average bottleneck's discharge flow can be achieved if queues are diminished upstream of the bottleneck; and 2) ramp metering can produce this outcome.

## **5. Conclusions**

This research has unveiled the cause and effect of time varying output flow through a freeway site that contains bottlenecks in series; and has confirmed that on-ramp metering can increase long-run average output flow and reduce system-wide delay. A horizontal curve becomes the site's downstream-most bottleneck. To combat the queues upstream of the curve bottleneck, a metering scheme was developed and field tested at the site. The two on-ramps (at Seamas and 43<sup>rd</sup> Aves) residing nearest upstream of the curve were metered by our control logic while other ramps operated as per Caltrans control logic. The major findings of this study are as follows:

- 1) A queue upstream of the horizontal curve diminished the curve's discharge flow. The queue tended to be more damaging, when it grew long, as it frequently did late in each morning rush. This is because the curve's long queue triggered the slowly progressing mechanism called the pinch effect by which i) a new bottleneck was pinched at the upstream weaving section near the 43<sup>rd</sup> interchange, and thereafter ii) this pinched bottleneck's severely low discharge flow starved the queue downstream of the weaving section. As a result of the pinch effect, the

- horizontal curve's discharge flow was fully recovered. This is good news. The bad news, however, is that the whole process entailed a long-lived (e.g. 6 minutes) reduction in the discharge flow from the horizontal curve.
- 2) Field experiments confirmed that in the early portion of each rush, higher long-run average discharge flows from the horizontal curve can be achieved by deploying restrictive (relaxed) metering at the Seamas on-ramp in response to the presence (absence) of short queues that frequently form upstream of the curve.
  - 3) Late in each rush, very restrictive metering deployed at the Seamas on-ramp could restore at least a portion of the horizontal curve's discharge flow. It seems that restricting inflows from the Seamas on-ramp reduced the frequency of disruptive vehicular lane-change maneuvers within the horizontal curve's queue downstream of the Seamas on-ramp.
  - 4) After a queue was dissipated upstream of the horizontal curve due to a pinch effect, very restrictive metering deployed both at Seamas and 43<sup>rd</sup> on-ramps was found to postpone another queue's growth at the horizontal curve. As a result, the site enjoyed long-lived (e.g. 10 minutes) high discharge flow from the horizontal curve before another long queue eventually formed there. Higher discharge flow was more prominent during periods in the rush when traffic demand was low.

The study highlights the benefits of devising freeway traffic control measures that are commensurate with whatever traffic demand level prevails. In case of higher demand (e.g. due to a strong economy), one might deploy metering schemes that combat freeway queues by coordinating control over numerous on-ramps, and not just at two ramps as was done in our field experiments.

In the coming months, we will further examine traffic flow details, behind the pinch effect. A greater understanding of the details could shed further light on what would constitute the most effective on-ramp metering schemes. We will keep Caltrans personnel informed of our findings in this regard.

## Acknowledgement

The Berkeley research team is much indebted to Caltrans personnel for their extraordinary collaboration and assistance on this project. In particular, we thank Jim Calkins, Brian Simi, Leo Anselmo, Gurdeep Sidhu, and Larry Hall from Caltrans District 3; and Melissa Clark and Zhongren Wang from Caltrans HQ.

## References

1. Ahn, S., Bertini, R.L., Auffray, B., Ross, J. H. and Eshel, O. (2007) Evaluating benefits of systemwide adaptive ramp-metering strategy in Portland, Oregon. *Transpn Res. Rec.* 2012, 47-56.
2. Allen, B.L. and Newell, G.F. (1976) Some issues related to metering or closing of freeway ramps. Part I. Control of a single ramp. Part II. Translationally symmetric corridor. *Transportation Science*, Vol. 10, No.3.
3. Athol, P.J. and Bullen, A.G.R. (1973) Multiple ramp control for a freeway bottleneck. *Highway Res. Rec.* 456, 50-54.
4. Banks, J.H. (1991) Two-capacity phenomenon at freeway bottlenecks: a basis for ramp Metering. *Transpn Res. Rec.* 1320, 83-90.
5. Bertini, R.L., Hansen S., and Bogenberger K. (2005) Empirical analysis of traffic sensor data surrounding a bottleneck on a German autobahn. *Transpn Res. Rec.* 1934, 97-107.
6. Cassidy, M.J. and Windover, J.R. (1995) Methodology for assessing dynamics of freeway traffic flow. *Transpn Res. Rec.* 1484, 73-79.
7. Cassidy, M.J. (2003) Freeway on-ramp metering, delay savings and the diverge bottleneck. *Transpn Res. Rec.* 1856, 1-5.
8. Cassidy, M.J. and Rudjanakanoknad, J. (2005) Increasing the capacity of an isolated merge by metering its on-ramp. *Transpn Res* 39B, 896-913.
9. Cassidy, M.J., Jang, K. and Daganzo, C.F. (2009) Smoothing effect on freeway bottlenecks: experimental verification and theoretical implications, *TRB 88<sup>th</sup> Annual Meeting*.
10. Chung, K, Rudjanakanoknad, J. and Cassidy, M.J. (2007) Relations between traffic density and capacity drop at three freeway bottlenecks. *Transpn Res* 41B, 82-95.

11. Daganzo, C.F. (1997) Fundamentals of transportation and traffic operations, Elsevier, New York.
12. Daganzo, C.F., Cassidy, M.J. and Bertini, R.L. (1999) Possible explanations of phase transitions in highway traffic. *Transpn Res* 33A, 365-379.
13. Daganzo, C.F. (2002) A behavior theory of multi-lane traffic flow. Part I: long homogenous freeway sections. Part II: Merges and the onset of congestion. *Transpn Res* 36B, 131-169.
14. Diakaki, C., Papageorgiou, M., and McLean, T. (2000) Integrated traffic-responsive urban corridor control strategy in Glasgow, Scotland - application and evaluation, *Transpn Res. Rec.* 1727, 101-111.
15. Edie, L.C. and Foote, R.S. (1960) Effects of shock waves on tunnel traffic flow. *Highway Research Board Proceeding*, Vol. 39, 492-505.
16. Haj-Salem, H. and Papageorgiou, M. (1995) Ramp metering impact on urban corridor traffic: field results. *Transpn Res* 29A, 303-319.
17. Hall, F.L. and Agyemang-Duah, K. (1991) Freeway capacity drop and the definition of Capacity. *Transpn Res. Rec.* 1320, 91-98.
18. Kerner, B.S. and Rehborn, H. (1996a) Experimental features of characteristics of traffic jams. *Physical Review E*, Vol. 53, No. 2, R1297-R1300.
19. Kerner, B.S. and Rehborn, H. (1996b) Experimental features of complexity in traffic flow. *Physical Review E*, Vol. 53, No. 5, R4276-R4278.
20. Kerner, B.S. (1999) Theory of congested traffic flow: self-organization without bottlenecks. *Transportation and Traffic Theory*, Pergamon-Elsevier, New York, 147-170.
21. Kerner, B.S. (2002a) Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks. *Physical Review E*, Vol. 65.
22. Kerner, B.S. (2002b) Empirical features of congested patterns at highway bottlenecks. *Transpn Res. Rec.* 1802, 145-154.
23. Laval, J.A. (2004) Hybrid Models of Traffic Flow: Impacts of Bounded Vehicle Accelerations. Ph.D. Dissertation, University of California, Berkeley.
24. Laval, J.A., Daganzo, C.F. (2006) Lane-changing in traffic streams, *Transpn Res* 40B, 251- 264.

25. Lighthill, M.J. and Whitham, G.B. (1955) On kinematic waves, Proceedings of royal society of London. Series A, Mathematical and Physical Sciences, Vol. 229, 317-345.
26. Lovell, D.J. (1997) Traffic Control on Metered Networks without Route Choice. Ph.D. Dissertation, University of California, Berkeley.
27. Minnesota Department of Transportation (2001) Twin cities ramp evaluation, Executive Summary, Cambridge Systematics, Inc., Cambridge, Mass.
28. Newell, G. F. (1993a) A simplified theory of kinematic waves in highway traffic, Part I: General theory. Transpn Res 27B, 281-287.
29. Newell, G. F. (1993b) A simplified theory of kinematic waves in highway traffic, Part II: Theory of Traffic Flow on Long Crowded Roads. Transpn Res 27B, 289-303.
30. Newman, L., Dunnett, A. and Meis, G. (1969) Freeway ramp control: what it can and cannot do, Los Angeles: California Division of Highways, District 7.
31. Papageorgiou, M., Haj-Salem, H. and Middelham F. (1998). ALINEA local ramp metering: Summary of field results. Transpn Res. Rec. 1603, 90-98.
32. Persaud, B.N., Yagar, S., and Brownlee. R. (1998) Exploration of the breakdown phenomenon in freeway traffic, Transpn Res. Rec. 1634, 64-69.
33. Richards, P.I. (1956) Shock waves on the highway. Operations Research. 4, 42-51.
34. Rudjanakanoknad J. (2005) Increasing freeway merge capacity through on-ramp metering. Ph.D. Dissertation, University of California, Berkeley.
35. Wattleworth, J.A. and Berry, D.S. (1965) Peak-period control of a freeway system-Some theoretical investigations. Highway Res. Rec. 89, 1-25.