A STUDY OF THE CAPACITY DROP PHENOMENON AT TIME-DEPENDENT AND TIME-INDEPENDENT BOTTLENECKS

Maha Salah El-Din El-Said El-Metwally

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Civil Engineering

Hesham A. Rakha, Chair
Ihab E. El-Shawarby
Antoine G. Hobeika

December 6th, 2010
Blacksburg, VA

Keywords: Capacity Drop, Onset of Congestion, Time-dependent Bottlenecks, Time-independent Bottlenecks

A STUDY OF THE CAPACITY DROP PHENOMENON AT TIME-DEPENDENT AND TIME-INDEPENDENT BOTTLENECKS

Maha Salah El-Din El-Said El-Metwally

ABSTRACT

The fact that traffic congestion upstream of a bottleneck causes a reduction in the discharge flow rate through the bottleneck has been well documented in several empirical studies. However, what has been missing is an understanding of the causes of these empirically observed flow reductions. An identification of these causes is important in order to develop various mitigation schemes through the use of emerging technology.

The concept of capacity drop can be introduced at time-independent bottlenecks (e.g. freeways) as well as time-dependent bottlenecks (e.g. signalized intersections). While to the author's knowledge no one has attempted to link these phenomena, the research presented in this thesis serves as a first step in doing so. The research uses the INTEGRATION simulation software, after demonstrating its validity against empirical data, to simulate time-independent and time-dependent bottlenecks in an attempt to characterize and understand the contributing factors to these flow reductions.

Initially, the INTEGRATION simulation software is validated by comparing its results to empirically observed traffic stream behavior. This thesis demonstrates that the discharge flow rate is reduced at stationary bottlenecks at the onset of congestion. These reductions at stationary bottlenecks are not recovered as the traffic stream propagates downstream. Furthermore, these reductions are not impacted by the level of vehicle acceleration. Alternatively, the drop in the discharge flow rate caused by time-dependent bottleneck is recoverable and is dependent on the level of acceleration. The difference in behavior is attributed to the fact that in the case of a stationary bottleneck the delay in vehicle headways exceeds the losses caused by vehicle accelerations and thus is not recoverable. In the case of vehicles discharging from a backward recovery wave the dominant factor is the delay caused by vehicle acceleration and this can be recuperated as the traffic stream travels downstream.

ACKNOWLEDGEMENTS

All my thanks go to Allah for all the bounties he has bestowed on me over my entire life.

I would like to express my gratitude and appreciation to my advisor, Prof. Hesham Rakha, who gave me the academic guidance and financial support to accomplish this thesis. During the course of my study Prof. Rakha was always very helpful and available whenever I needed him. His sincere motivation and continuous care make him like an older brother for me.

I am also very thankful to Dr. Ihab El-Shawarby, my committee member, for his continuous advice and invaluable help during my research. I also would like to deeply thank Dr. Antoine Hobeika, my committee member and also my professor for two graduate classes, for his support and encouragement.

I would like to thank my colleagues at Virginia Tech for their cooperation and for the good time we shared together during this research period.

Finally, I would like to thank my family, my father, my mother and my sisters for their continuous care and support. I want especially to thank my father for his encouragement for me to pursue my M.S. degree. I would like to dedicate my thesis to my husband, Ahmed, and thank him for his loving support and sincere encouragement in every step through our life together; he is always there for me.

ATTRIBUTIONS

My thesis advisor, Prof. Hesham Rakha, aided in the writing and organization of the research presented in this thesis. In addition, every chapter of this thesis was extensively guided by him. Besides, Prof. Rakha is the co-author on all the papers included in this thesis.

Edward Chamberlayne is the first author of the paper presented in chapter 3. In this chapter, he conducted many simulation runs and wrote the first draft of many parts in this paper. In addition, he is cited as a third author on the paper in chapter 4, where he shared some ideas during the working meetings for this paper with Prof. Rakha and me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 : INTRODUCTION

1.1 PROBLEM OVERVIEW

The drop in highway capacity at the onset of congestion is a very important phenomenon from the perspective of traffic management because it suggests that one can increase the transportation system throughput by ensuring that congestion does not occur. This reduction in capacity will result in increased travel times, excessive delays, and increased fuel consumption rates [1]. What still remains an issue of debate is the causes and influencing factors that affect this capacity reduction. Understanding the cause of the capacity drop and quantifying its magnitude is important in estimating the capacity of a highway system [2] and identifying means to mitigate these losses. Such answers would facilitate in reducing the negative impacts caused by recurring congestion. Not only is the capacity drop important operation-wise, but it is also important from a safety perspective. It was reported that the crash frequency is 300% higher in congestion than in uncongested conditions at similar flows [3].

In general, the capacity drop phenomenon can be observed on uninterrupted flow facilities (i.e. freeways) as well as interrupted flow facilities (e.g. signalized intersections). The main difference between both cases is that on freeways, congestion is typically discharged from a stationary bottleneck while at signalized intersections, vehicles discharge from a backward recovery shockwave. In other words, the queue discharge point at a freeway bottleneck is fixed, as each vehicle starts accelerating from the same location within the bottleneck. On the other hand, the queue discharge point at signalized intersections is a backward moving point where each vehicle starts accelerating from its position within the queue. Hence, further upstream vehicles reach the stop line with higher speeds than the vehicles closer to the stop line.

A number of important questions need to be addressed, as follows. (a) How can the capacity drop be replicated using an agent-based approach? (b) What factors cause the capacity drop? (c) Are the losses caused by the capacity drop permanent or are they temporary recoverable losses?

1.2 THESIS OBJECTIVES AND CONTRIBUTIONS

In light of the above discussion, the research presented in this thesis attempts to address three objectives. The first objective entails validating the INTEGRATION agent-based approach in replicating field observed behavior. The second objective is to investigate the impact of different factors on the capacity drop. The final objective is to compare the drop in the discharge flow rates for stationary (time-

independent) bottlenecks on freeways to moving (time-dependent) bottlenecks at signalized intersections, by quantifying the spatial and temporal aspects of these losses.

The contributions of the thesis can be summarized as follows:

1. The research demonstrates the validity of the INTEGRATION simulation software in producing a reduction in the discharge flow after the onset of congestion and thus replicating empirically observed traffic stream behavior.
2. The study demonstrates that reductions in discharge flow rates can be replicated through the accurate modeling of driver acceleration behavior without the need to introduce a discontinuity in the steady–state fundamental diagram.
3. The study identifies and quantifies the impact of different underlying factors (e.g. acceleration level, type of bottleneck, total demand, and percentage of heavy vehicles) on the discharge flow rates from different bottlenecks.
4. The study demonstrates that flow reductions at stationary (time-independent) bottlenecks produce losses that are not recovered downstream, whereas at moving (time-dependent) bottlenecks flow reductions are found to be spatially and temporally confined.

1.3 RESEARCH APPROACH

In order to fulfill the thesis objectives, the concept of the capacity drop is first discussed by reviewing the relevant state-of-the-art literature on the topic. This investigation analyzes different congestion conditions at freeway bottlenecks. The approach is then applied to signalized intersections by analyzing the start loss associated with the discharge of stopped vehicles after the signal turns from red to green. The research uses the agent-based INTEGRATION simulation software for the analysis. Consequently, a validation of the INTEGRATION software is performed by modeling and replicating a number of bottleneck scenarios reported in the literature. After validating the INTEGRATION software, the impact of different factors on the discharge flow rate is investigated. This is done using the INTEGRATION software as it is difficult to determine the influence and the contribution of different factors through empirical studies. Furthermore, simulation runs are conducted in order to investigate the differences in the nature of the discharge flow rates at time-dependent and time-independent bottlenecks.

1.4 THESIS LAYOUT

This thesis is divided into five chapters. The first chapter provides a brief introduction to the problem, which is followed by chapter two that provides an in-depth review of the relevant literature discussing the issue of capacity drop at uninterrupted flow facilities (i.e. freeways) as well as the interrupted flow facilities (e.g. signalized intersections). Subsequently, chapter three presents a paper on a simulation study of freeway bottleneck flow reductions at the onset of congestion. The fourth chapter presents a paper comparing queue discharge rates from time-dependent and time-independent bottlenecks. Finally, chapter five presents the thesis conclusions and recommendations for further research.

CHAPTER 2 : LITERATURE REVIEW

2.1 INTRODUCTION

As was mentioned earlier, the research presented in this thesis attempts to model and characterizes the reductions in discharge flows at the onset of congestion. This objective is achieved first by investigating the relevant literature by starting first with explaining the definition of capacity and accordingly the definition of capacity drop. Hence, different perspectives from previous research studies discussing the capacity drop on uninterrupted flow facilities (i.e. freeways) is presented. Thereafter, the capacity drop issue on interrupted flow facilities (e.g. signalized intersections) is introduced as well.

2.2 CAPACITY DROP PHENOMENON

2.2.1 The Concept of Capacity

The concept of capacity is one of the most debatable issues in the field of traffic flow theory, many decades ago up to this date. Those debates about this issue not only are related to the numerical value of the capacity on different transportation facilities, but also extend to the notion of the capacity drop. However in order to understand those debates about the concept, we should first define capacity. The 2000 Highway Capacity Manual (HCM) [4] defines capacity as:

"The maximum sustainable flow rate at which vehicles or persons reasonably can be expected to traverse a point or uniform segment of a lane or roadway during a specified time period under given roadway, geometric, traffic, environmental, and control conditions; usually expressed as vehicles per hour, passenger cars per hour, or persons per hour".

Furthermore, it is widely accepted that the concept of capacity is presented throughout the speed-flow relationship curves. This relationship has a parabolic shape, which consists of two branches, as shown in Figure 1. The upper branch represents the uncongested flow that begins with the free flow speed at very small flow rates. Subsequently, as the flow rate increases the speed decreases until the capacity is reached at the vertex of the parabola. On the other hand, the lower branch represents the congested flow. This branch begins when the vehicles start flowing from a zero speed and accelerating until reaching the capacity.

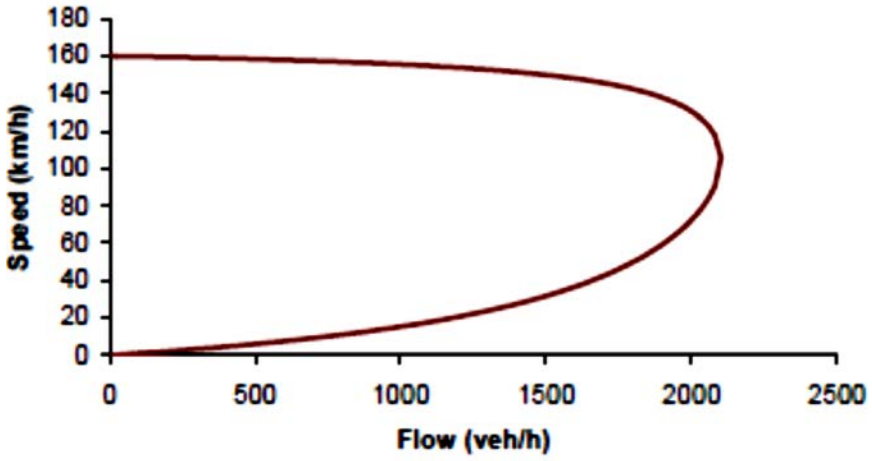


Figure 1: Capacity on Continuous Speed-Flow Relationship [5]

Thus, under the current HCM definition, capacity is considered to be a "deterministic" value that defines the point above which the operations at a facility will "breakdown." Athol and Bullen [6] were among the first researchers who suggested that the probability of a transition from uncongested to congested flow depends on the value of flow and that the expected time to flow breakdown is a declining function of flow. Hall and Agyemang-Duah [7] also raised important arguments about the validity of the definition of freeway capacity. Based on empirical traffic data from the Queen Elizabeth Way (QEW) in Mississauga, west of Toronto, Ontario, Hall and Agyemang-Duah concluded that the mean value of the estimated normal distribution of maximum flows at a single location may not be appropriate to describe "capacity" at a specific location.

Recent studies have provided additional empirical evidence that under similar environmental conditions, different capacities can be observed on a single section of a freeway. According to Elefteriadou, et al. [8], existing "*models assumed that breakdown would occur at capacity flows only, and hence capacity could be measured in the field just prior to breakdown*." To address this shortcoming, Elefteriadou, et al. [8] collected and analyzed traffic data at various ramp merge interchanges. The main conclusion of this study was the empirical observation that "*breakdown was not the direct result of peak volumes, and breakdown often occurred at flows lower than the observed capacity*." They found that the typical sequence of events prior to breakdown were: (a) entrance of a large cluster to the freeway stream; (b) subsequent small speed drop at the beginning of the bottleneck; and (c) spread of the queue upstream and subsequent flow breakdown.

In a later study, Persaud, et al. [3] defined flow breakdown as a "*description of freeway operation near a bottleneck entrance during a period when there is a change from operation with*

vehicles flowing freely to operation with a queue present." They used the Highway 401 and Gardiner Expressway data to show that there is a relationship between freeway flow and probability of breakdown.

Later, Lorenz and Elefteriadou [9] used data from two sites on Highway 401 in Toronto, Canada. They concluded one more time that the capacity of freeways cannot be defined "*by a single value – specifically a single rate of flow*." They also concluded that the implication in the HCM that states that breakdown is expected to occur on the freeway once the flow rate corresponding to capacity is reached or exceeded for the specified time period cannot be supported on the basis of empirical data analysis.

In summary, traffic breakdown called the "breakdown phenomenon" occurs mostly at freeway bottlenecks, in particular on-ramp bottlenecks. The traffic breakdown is a probabilistic phenomenon that can occur over a range of flow rates. At the same on-ramp bottleneck, traffic breakdown is observed at different flow rates in different days. The probability of breakdown is an increasing function of flow rate downstream of the bottleneck and the length of the time interval considered.

2.2.2 The Concept of Capacity Drop

One of the most important unresolved issues about capacity is the issue of the capacity drop. This issue can be illustrated as shown in Figure 2. The figure presents a discontinuous speed-flow relationship that is different from the one presented in Figure 1. Based on the continuous relationship in Figure 1, capacity does not change if approached from stable "uncongested" flow or approached from unstable "congested" flow. However, the relationship in Figure 2 is a discontinuous relationship, where capacity is higher when approached from stable flow than when approached from unstable flow. This is what is being called capacity drop that is assumed to take place at the onset of congestion.

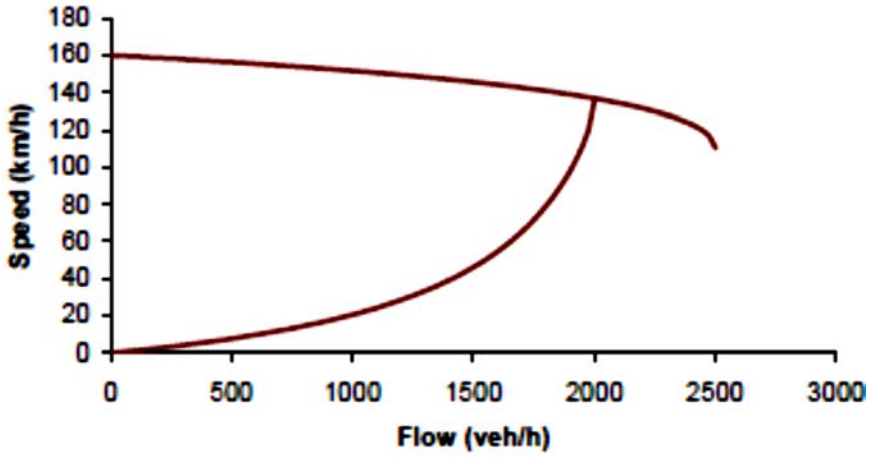


Figure 2: Capacity Drop on Discontinuous Speed-Flow Relationship [5]

In order to be able to investigate the capacity drop phenomenon, the concept of congestion needs to be presented in general. The general mechanism of the occurrence of congestion can be illustrated throughout the detailed investigation of the discontinuous speed-flow relationship, presented in Figure 3. As can be interpreted from the figure, the curve is divided into four regions; regions 1 and 2 belong to the stable flow portion of the curve, whereas region 3 and 4 belong to the unstable flow portion [10].

Throughout region 1, driver's speed, beginning from the free-flow speed, remains almost constant to relatively high flow rates. In response to increasing flow rates, speed begins to decline until reaching the speed at capacity, as shown by region 2. Once demand exceeds capacity, breakdown or congestion occurs and queue propagates upstream of the point of breakdown. Region 4 shows the unstable operating conditions within the queue, whereas region 3 reflects the queue discharge flow. In this region, vehicles discharge from a queue into uncongested downstream, where the flow remains constant, while speed increases as drivers accelerate after departing the queue [10].

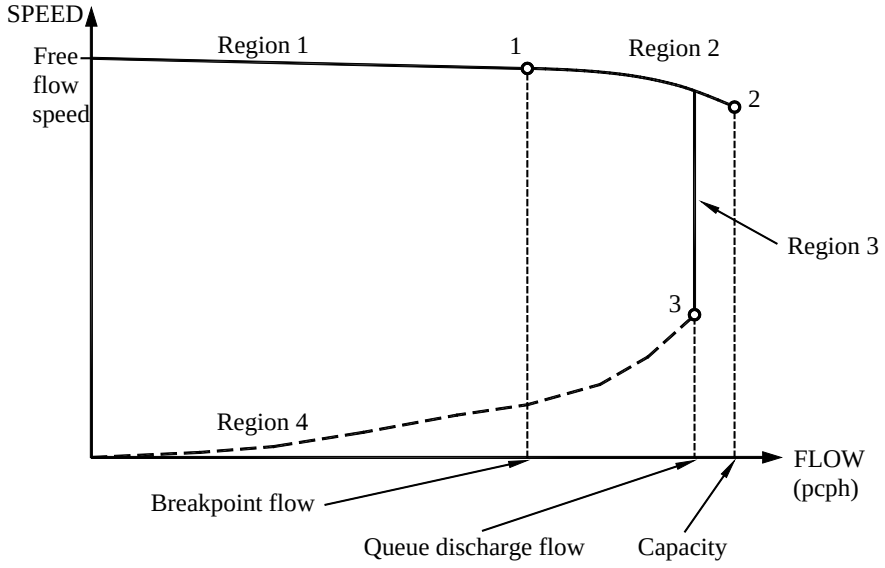


Figure 3: Mechanism of Congestion on Discontinuous Speed-Flow Relationship [10]

2.3 CAPACITY DROP AT DIFFERENT FACILITY TYPES

In general, the phenomenon of capacity drop can be observed on uninterrupted flow facilities (i.e. freeways) as well as the interrupted flow facilities (e.g. signalized intersections). At uninterrupted flow facilities, the congestion takes place at the activation of stationary bottlenecks, where queue starts to form upstream the bottleneck. Activation of the bottleneck means that the flow reduction is not an external effect caused by a drop in demand or by a queue formed from another downstream bottleneck

[1]. In order to investigate the claim that capacity diminishes once a queue forms, capacity before congestion is compared with queue discharge flow. On the other hand, at interrupted flow facilities, the congestion occurs once the traffic signal indication turns red and the queue starts to form upstream the intersection stop line. Once the signal turns green, the resulting queue starts to discharge. The vehicles move sequentially, where each vehicle is being delayed until the vehicle ahead of it starts moving. The time until each successive vehicle begins to move is known as the start-up lost time [11], which is most for the lead vehicle and diminishes for the following vehicles until the headway between the vehicles becomes constant.

Generally, the concept of capacity drop on both facility types; uninterrupted and interrupted flow facilities, is the same: the onset of congestion results in a drop in the roadway capacity. Nevertheless, the main difference between both cases is that on freeways congestion occurs at stationary bottlenecks at a time independent bottleneck. While at signalized intersections the bottleneck is time dependent where congestion occurs at the onset of red. Besides, the queue discharge point in freeways is fixed, as each vehicle starts accelerating from the queue speed at approximately the same location at the bottleneck. On the other hand, queues at signalized intersections discharge from a backward recovery shockwave where each vehicle starts accelerating from a different point. Hence, further upstream vehicles reach the stop line at higher speeds.

2.3.1 Capacity Drop on Uninterrupted Flow Facilities

The concept of two capacities or the “capacity drop” was first introduced by Edie in 1961 [12] where congested and uncongested flows in the Lincoln Tunnel in New York were analyzed. This empirical study provided initial support to the concept of flow reductions downstream of a congested bottleneck. As mentioned earlier, the issue of capacity drop is very debatable among different capacity analysis studies. Among those studies concerning capacity analysis on freeways, some studies have reported a drop in the maximum flow once a queue forms just upstream a bottleneck, whilst others have indicated that no such drop occurs. The onset of congestion on freeways, as previously illustrated, is followed by queue forming upstream the stationary bottleneck. This bottleneck is considered active whenever the demand exceeds the capacity due to, for example, a reduction in number of travel lanes, on-ramp merging into the freeway, or horizontal curve on the freeway. In presenting the arguments related to freeway capacity drop, some of the studies investigating the reduction in queue discharge flows just upstream of the bottleneck are presented and synthesized hereinafter.

First, it should be investigated if such a capacity drop can be related to other measurements or not. Chung et al inferred that increase in the freeway density is always followed by a drop in capacity [1]. Three different types of bottlenecks on three different freeway segments were investigated to examine the relation between the vehicular density and the drop in capacity. One was produced by an on-ramp merge, whereas the second was produced by a reduction in the number of travel lanes, and the last by a horizontal curve on the freeway. Using video cameras, traffic data were collected and analyzed at the three bottlenecks, when they were verified to be active; i.e. the reduction in flow is the bottleneck's internal feature that is called capacity drop and not an external effect, as mentioned earlier [1]. As an example illustration, the analysis of the on-ramp merge bottleneck is presented here to demonstrate the analysis procedure. The merge bottleneck exists on the freeway stretch, shown in Figure 4, of the north-bound Interstate 805 in San Diego, CA. The bottleneck occurs between X_2 and X_3 due to the merge of the on-ramp at 47^{th} St/Palm Ave with the freeway [1]. The virtual cumulative arrivals at locations X_1 through X_4 and the corresponding density are plotted against time in Figure 5 (a) and (b), respectively. As shown from the figure, the arrival curves at X_2 and X_3 diverge at about t = 6:15 to reveal an intervening queue. While, the curves X_3 and X_4 are superimposed continuous as verification that the bottleneck is an active bottleneck (i.e. not affected by downstream bottlenecks). The dotted trend lines indicate a reduction in flow that begins at 6:23 of approximately 10%. The mechanism of this bottleneck's capacity drop began by formation of a queue in the freeway shoulder lane until the accumulation of this queue exceeds a critical number that is accompanied by the occurrence of a flow reduction [1]. By measuring the vehicle accumulation across all lanes between X_1 and X_3 and dividing this number by the distance between the two locations, the density was obtained. As shown in Figure 5 (b), the densities versus time on the day of the activation of the bottleneck are presented. The density rose gradually beginning at t = 6:15 until reaching a density, which is assumed to be the critical density, that corresponds the capacity drop. Moreover, the concurrence between this critical density and the capacity drop was reproducible across days to a reasonable extent.

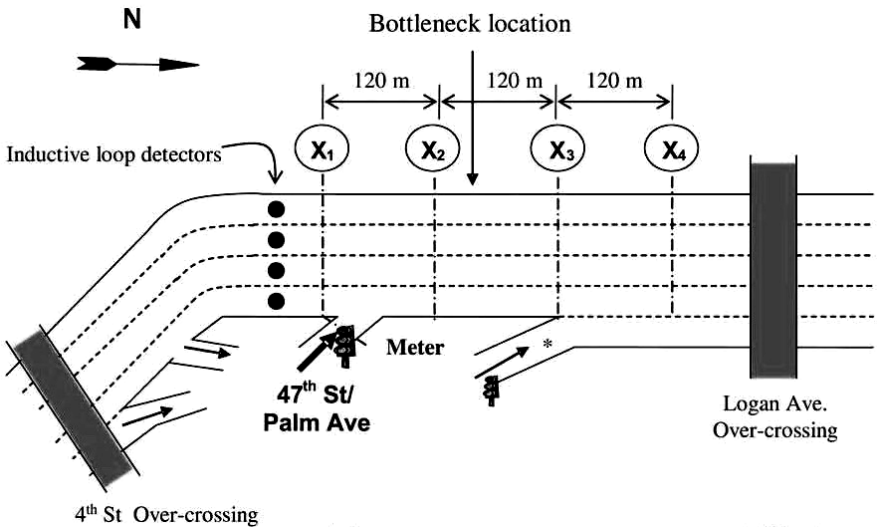


Figure 4: Layout of On-ramp Merge Bottleneck in San Diego, CA [1]

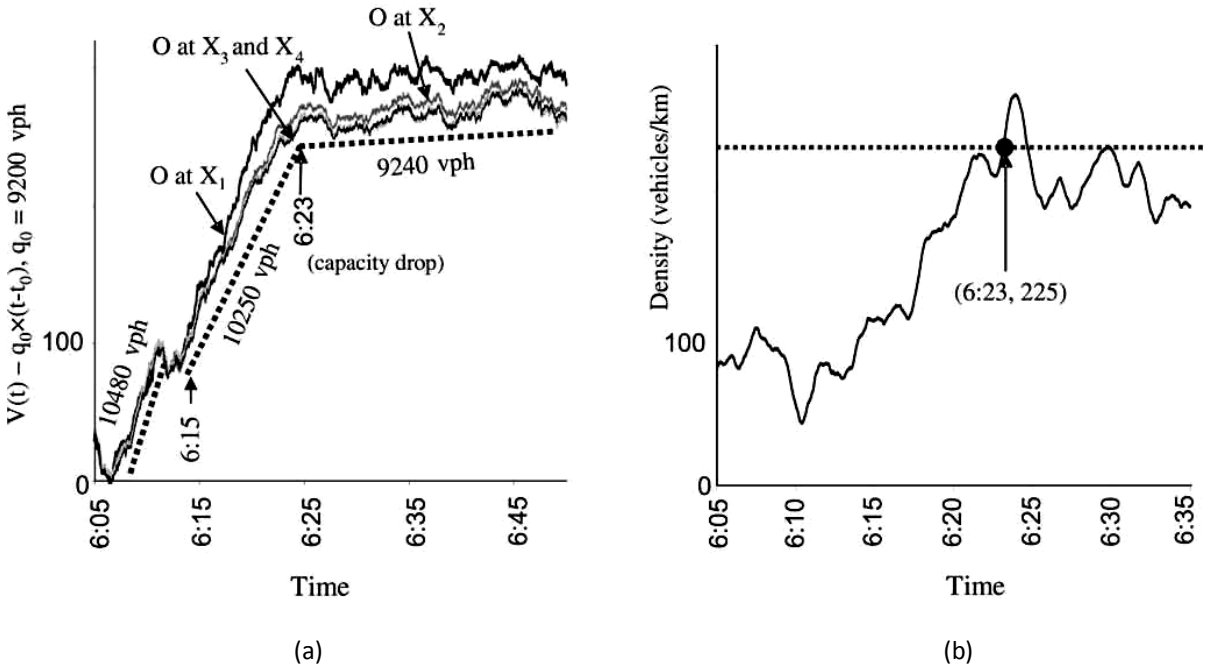
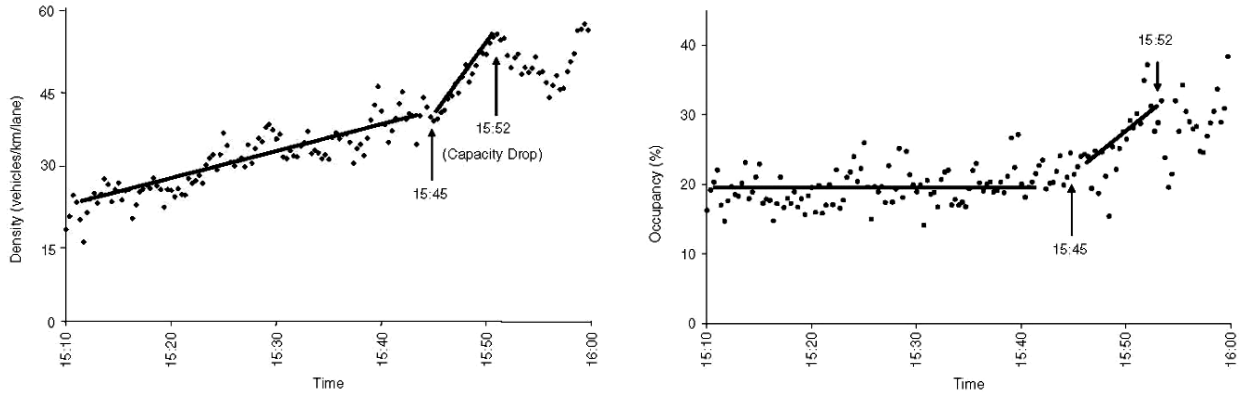


Figure 5: Cumulative Plots (a) and Density Distribution (b) at the On-ramp bottleneck [1]

Accordingly, it was concluded that the increase in density in all lanes near the bottleneck correlates with the capacity drop [1]. However, as the three bottlenecks had different number of travel lanes, the density corresponding to the capacity drop was normalized by the number of lanes and averaged across days at each site to obtain the critical density. The normalized critical density for the merge bottleneck matches the normalized critical density of the horizontal curve bottleneck, which was

found to be approximately 55 veh/km/lane. The critical density at the lane reduction bottleneck was found to be smaller by approximately 10 veh/km/lane. According to the authors, this difference may be due to errors in the data extraction procedures [1]. Furthermore, it was acknowledged that occupancy also relates to capacity drop. However, the authors concluded that real-time control schemes that respond to occupancy may not be as good given that the occupancy was less sensitive to the changes in traffic stream, as demonstrated in Figure 6. It can be illustrated from Figure 6 (a) that there is a gradual rise in density starting earlier than the obvious rise some minutes prior to the capacity drop, where these changes were measured as they occurred. Accordingly, the control actions will be based on these gradual changes. On the other hand, Figure 6 (b) shows that the gradual rise in the density is not reflected in the occupancy distribution. Hence, an occupancy-based scheme would in this case have less time to respond to a capacity drop [1].



(a) (b)

Figure 6: Density (a) and Occupancy (b) Distribution at the Curve Bottleneck [1]

Nevertheless, what was concluded in the previous study [1] that a specific value for critical density results in a capacity drop is somewhat questionable. Generalizing a fixed critical density value to be applied at any freeway could not be accepted unconditionally, as the capacity drop can result from factors other than increased density. Contrarily, increased traffic density can occur without a corresponding drop in capacity. For instance, a sudden lane changing decision from one driver can block a lane and increase its density, while the traffic volume is not high or near capacity. This discussion implies that not only a high traffic density can trigger a capacity drop, but it should be accompanied by high flow rates. In spite of that, no deterministic value for density of flow can accompany the reduction in capacity. On the contrary, the occurrence of breakdown and the corresponding capacity drop is a

more probabilistic event. An earlier study in 1998 by Persuad et al [3] stated that the occurrence of breakdown has a probabilistic nature. To explore the probability of breakdown traffic data were collected at three different sites in Toronto, Canada using surveillance cameras and automatic collection of speed and flow data from loop detectors. The study found a capacity drop ranging from 11% to 17% in two of the sites, while in the third site the capacity drop was 26%. This increased value in the third site was considered but no definite conclusions were made and further work was needed [3]. Furthermore, the probability of breakdown at various traffic flow levels was investigated. As shown from Figure 7, the probability of breakdown based on flows for the three sites was plotted. In general for the three sites, the probability of breakdown rises by increasing the flow rates. It was found that by maintaining the pre-queue flows at the same level as those that occur after a queue forms, then the probability of breakdown is almost negligible. By increasing these flows 20% above the queue discharge flow, the probability of breakdown is only 10%. For more increased flows the probability of breakdown rises dramatically. Hence, it was suggested not to operate at pre-queue exceeding the queue discharge more than 20% [3].

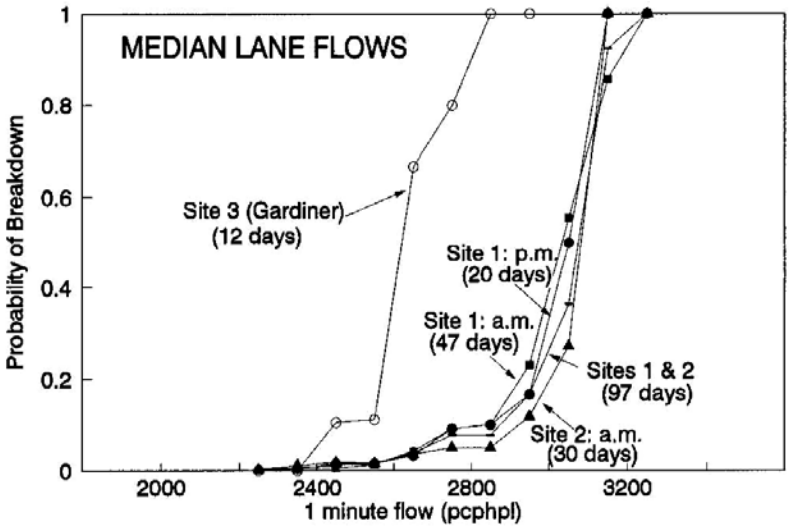
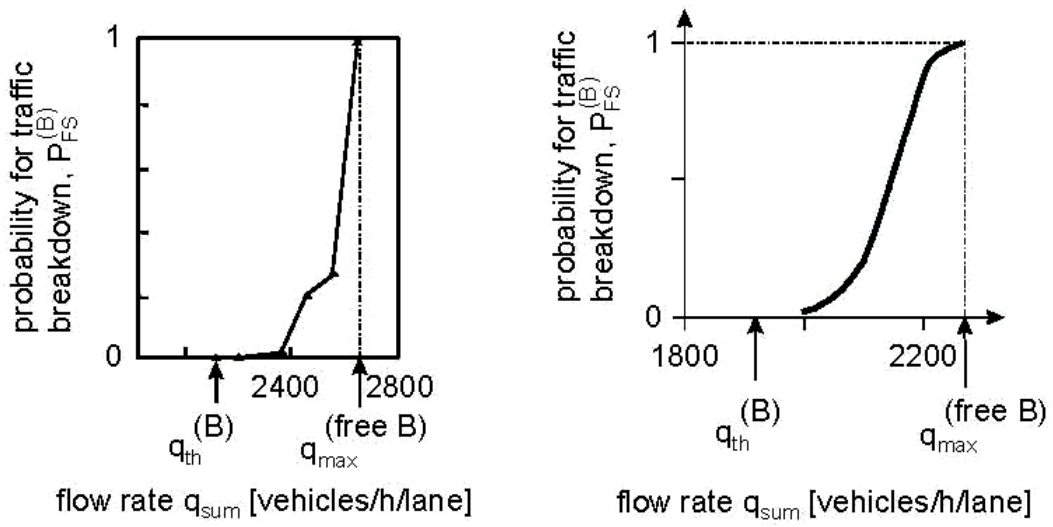


Figure 7: Probability of breakdown based on flows for three Sites in Toronto, Canada [3]

Furthermore, in more than one study by Kerner, e.g. [13, 14], that built his hypothesis based on the findings of Persuad [3], Kerner investigated the breakdown phenomenon throughout his proposed three-phase traffic theory [15]. In his theory, Kerner represented the occurrence of breakdown by the transition from the free flow phase (F) to synchronized flow phase (S). As shown in Figure 8, Kerner [14] agrees that the breakdown is a probabilistic phenomenon , and not deterministic as concluded by Chung et al [1].



(a) (b)

Figure 8: Measured (a) and Theoretical (b) Breakdown Probability [14]

With respect to a debate on the existence of capacity drop, in 1990 Hall and Hall [16] investigated the effects of formation of an upstream queue on speed and flow. Data were collected from the Queen Elizabeth Way in Ontario, Canada. It was claimed, by the authors [16], that there is no reduction in capacity at bottlenecks downstream of a queue and that upstream queue formation had no effect on flow rates. However, this queue formation caused a reduction in observed speed [16].

Nevertheless, in 1991, Hall and Agyemang-Duah [7] came back and denied this claim, saying that Hall and Hall [16] did not have as good information with which to ascertain the time of the queue. In addition, the flows in the data they used were not heavy enough. However, in the 1991 study [7], they concluded that once queue forms upstream of the bottleneck a capacity drop appears as a consequence of the way drivers accelerate away from the queue. Their conclusion was based on data collected from the same place, shown in Figure 9, like the previous study in Queen Elizabeth Way in Ontario, Canada [16] and the results show a capacity drop of about 5 to 6% [7]. In addition to the idea of the capacity drop within the bottleneck, the paper argued that the location of a possible reduction in capacity should be identified [7]. The authors stated that other studies measure the capacity before and after the congestion at the location of the bottleneck, as station 22 in Figure 9. Measuring the flows at this location yields a two branched occupancy-flow relationship as shown in Figure 10 (a). The left branch represents the stable (uncongested) flow, which is higher than the right branch, which represents the unstable (congested) flow. The authors argued that this station is not operating at capacity during

unstable flow [7]. Consequently, the discontinuity of the curve does not indicate a capacity drop. However, they concluded that the right place to measure capacity drop is at station 25. Measuring at this location yields a one branch occupancy-flow relationship as shown in Figure 10 (b). This branch consists of pre-queue flows and queue discharge flows. The difference between these two curves presents the capacity drop [7]. In this study, a truck percentage of 6 percent was reported in the traffic stream [7]. Nevertheless, the location of station 25 is still not convincing enough. This is because of the existence of another on-ramp merge from Dixie Rd., whose impact is to be considered at this location measurements. In addition, station 25 is far downstream from the bottleneck location at station 22, which assumes that the bottleneck impact extends to that far downstream.

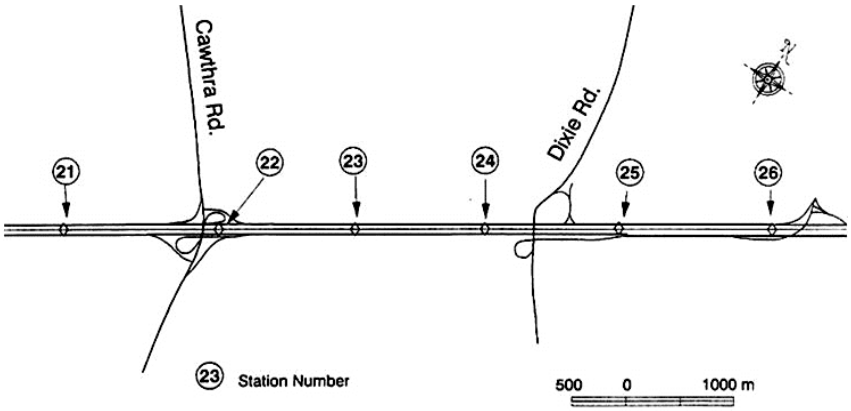
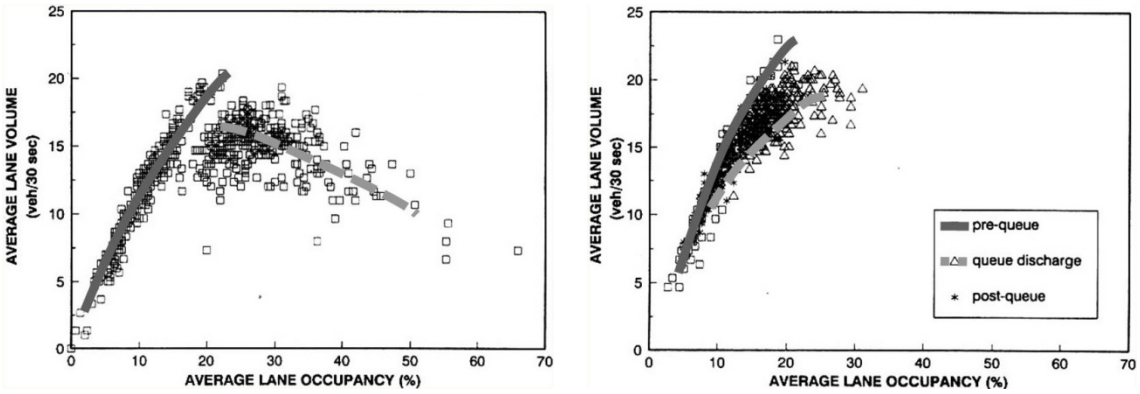


Figure 9: Layout of Bottleneck Location on Queen Elizabeth Way in Ontario, Canada [7]



(a) (b)

Figure 10: Occupancy-Flow Relationships at (a) Station 22 and (b) Station 25 [7]

Cassidy and Bertini 1999 [17] studied the capacity drop phenomenon along the QEW as was done in the previous study [7]. A capacity drop of 8 to 9 percent was found along the QEW, as compared

to the 5 to 6 percent drop found in [7], by visually comparing transformed cumulative curves and cumulative occupancy against time created by data from downstream loop detectors [17].

One more argument about the issue of capacity drop is raised in one study by Banks [18]. Even though the study concluded that bottleneck capacities decrease when flow breaks down, Banks concluded that this applies to the capacities of individual lanes [18]. Four different bottlenecks in four sites in San Diego in California were investigated. As shown in Figure 11 (a), it was found that capacity decreased after breakdown by approximately 10% in the left lane for Site 1. However, the capacity when averaged across all lanes decreases only 3%, as in Figure 11 (b). The left lane in the other three sites had a capacity drop of approximately 4.5% in two sites and 0.6% in the third. On the other hand, flow averaged across all lanes does not show any decrease in the three sites [18].

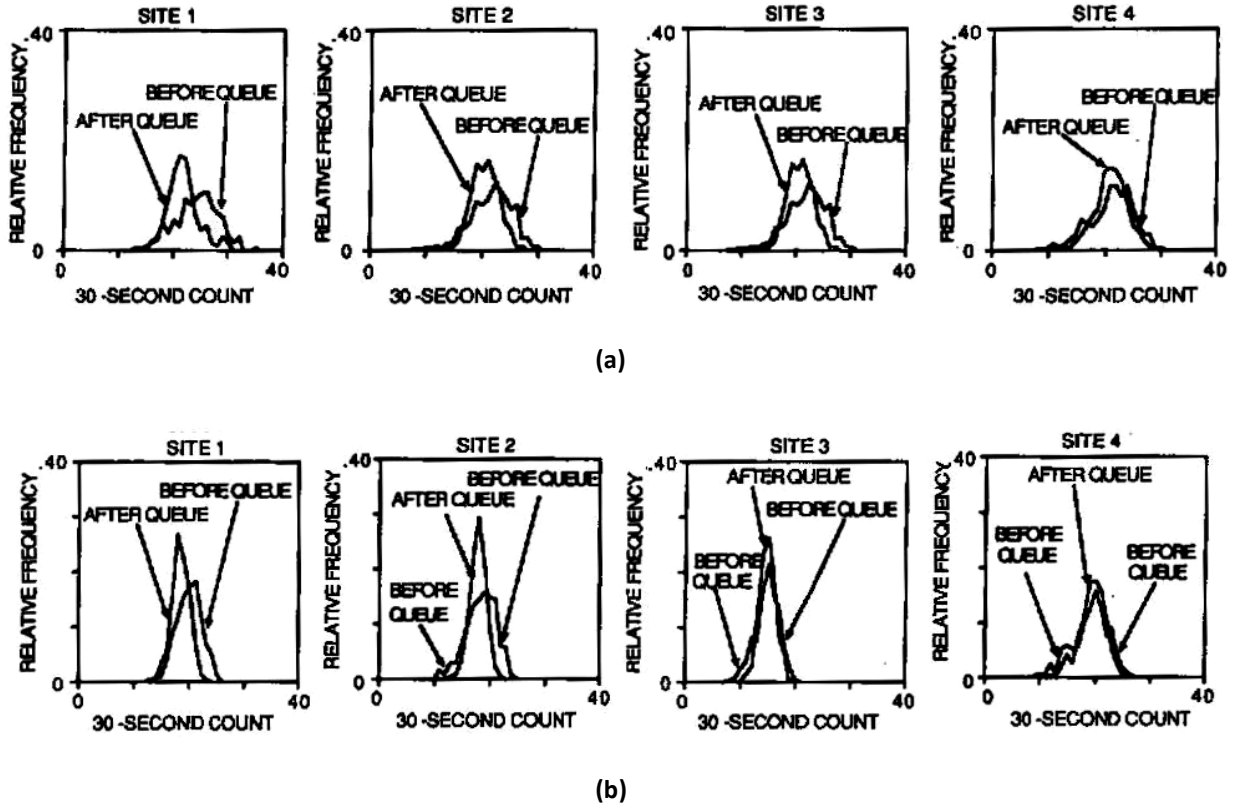


Figure 11: Flow Frequencies Before and After Queue for (a) Left Lane and (b) All Lanes at 4 Sites in San Diego, CA [18]

Another study carried out by Papageorgiou et al. [19] reported that the measured downstream flow from merged bottleneck was lower than the capacity flow by 5-20% in different investigations [19]. In this study it was stated that the capacity drop is a result of vehicle acceleration at the bottleneck. In

addition different driver attitudes and vehicle capabilities result in longer gaps between the accelerating vehicles resulting in the capacity drop [19]. Given that one explanation for the reduction in discharge flow after the onset of congestion relates to driver acceleration levels, another study by Snare in 2002 found that drivers normally accelerate at 60 percent of the maximum acceleration level of their vehicle; *"Driver factors were observed that fit a normal distribution with a mean of 0.6 and a standard deviation of 0.08"* [20]. This confirmed a previous guideline published in 1954 by AASHO [21]. This conclusion is important since the flow reductions are influenced by how vehicles accelerate downstream of the bottleneck.

2.3.2 Capacity Drop on Interrupted Flow Facilities

As capacity on uninterrupted flow facilities (freeways) was discussed in the previous section, capacity on interrupted flow facilities (signalized intersections) will be investigated, assuming that the green signal is available all the time. This is known as the saturation flow rate, which is defined in the 2000 HCM [4] as:

"The equivalent hourly rate at which previously queued vehicles can traverse an intersection approach under prevailing conditions, assuming that the green signal is available at all times and no lost times are experienced, in vehicles per hour or vehicles per hour per lane."

Furthermore, the saturation flow rate represents the capacity at signalized intersections weighted by the ratio of the cycle length to the effective green time. Nevertheless, saturation flow rate at signalized intersections is analogous to the concept of capacity on freeways, as both consider 100% green all the time. In practice, the saturation flow rate is not measured directly from field observations. However, it can be obtained using its relation with the saturation headway. The headway in general is the time in seconds between two successive vehicles passing the stop line, measured from the same reference points (e.g. front axle to front axle). While, the saturation headway is the stable headway between vehicles occurring after the impact of the start-up lost time vanishes. Having the saturation headway measured from field observations, the saturation flow rate is calculated as the inverse of the saturation flow headway in the following equation.

$$s = \frac{3600}{h}$$

As stated before, the headway for the first few vehicles is affected by the start-up lost time that induces increases in their headways. This increase is attributed to the driver perception-reaction time and the acceleration time. This increase is longest for the first vehicle in the queue and diminishes for

the following vehicles successively. The HCM assumes that the start-up lost time diminishes after the fourth vehicle and a stable headway continues after that [4]. Figure 12 shows a conceptual plot of the distribution of headways for different queue positions, based on the 2000 HCM. It can be concluded from the figure that the first four vehicles have greater headways, while starting from the fifth vehicle the stable saturation headway is achieved [4].

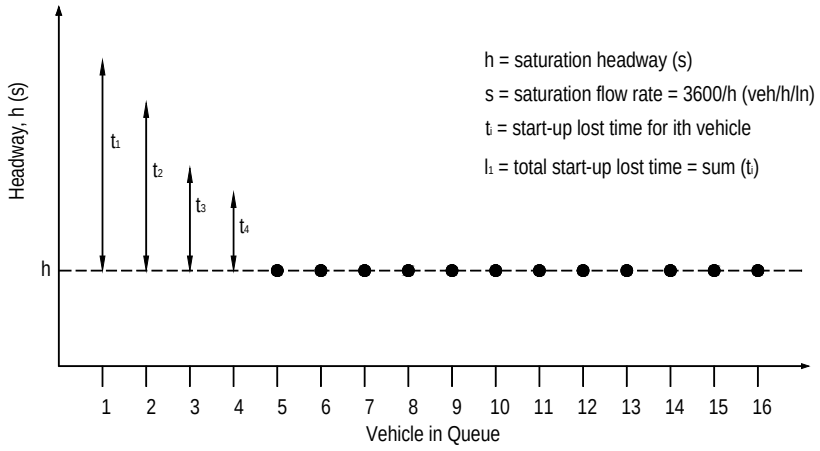


Figure 12: HCM Headway Distribution with Queue Position [4]

Using the relationship between saturation flow rate and saturation headway presented earlier, the above distribution of the queue position headways can simply be translated into a distribution for the saturation flow rates, as presented in Figure 13. In agreement with the headway distribution, the saturation flow rate for the first vehicle is the smallest. However, it increases within the first four vehicles until reaching a constant saturation headway starting from the fifth vehicle in the queue. The saturation flow distribution implies that the ending point of start-up lost time, which is most for the first vehicle and diminishes for the following vehicles successively, is the beginning point for achieving a stable saturation flow rate.

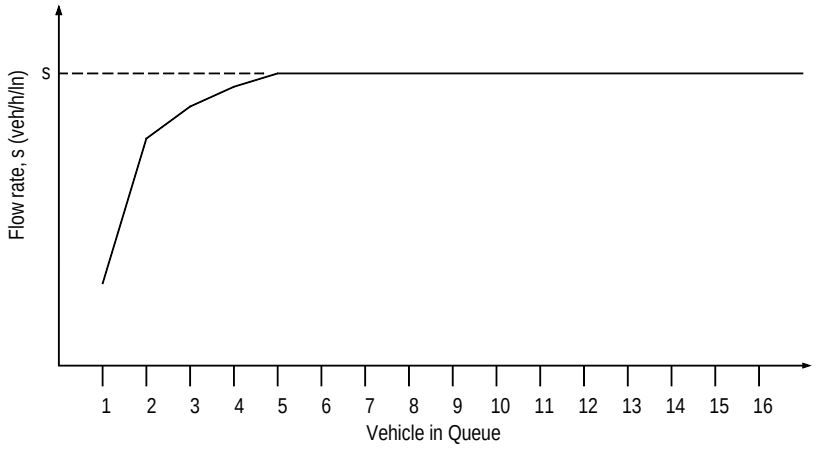
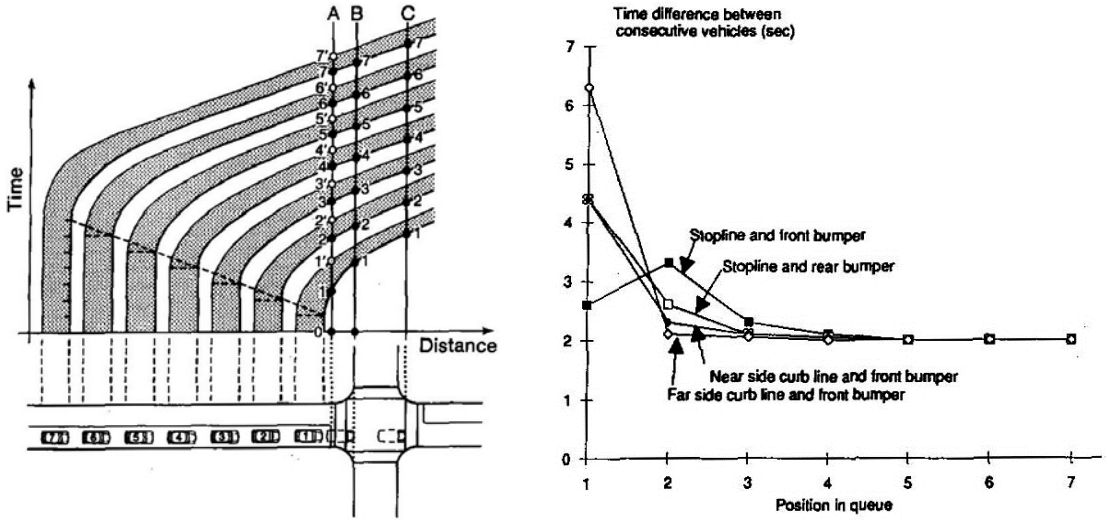


Figure 13: Saturation Flow Rates Corresponding to the HCM Headway

However earlier, Teply and Jones [22] argued that although the general descriptions of saturation flow rates based on the HCM are similar to other guidelines (e.g. [23, 24]), this applies only for close-to-ideal intersections. All these guidelines recommend the use of measured saturation flows rather than the default values provided in these documents. They all agree on the great variability in saturation flows caused by different urban conditions and various geometric and traffic configurations. Although there are many similarities in the saturation flow adjustments recommended in these documents, actual procedures for each of them differ [22]. Example of the different measurement procedures are presented in Figure 14 showing various measurement points for queue discharge headways, identified in Figure 14 (a) as lines A, B and C. Variation in the headways caused by the different combinations of the different reference points with the different measurement locations (e.g. stop line + front bumper, nearside curb + front bumper) are presented in Figure 14 (b). Two basic shapes could be recognized, where one shape represents the time difference needed by the first vehicle is the greatest applies for three combinations. While the other shape represents the greatest time difference occurs for the second vehicle [22].



(a) Headway measurement locations (b) Headway distribution

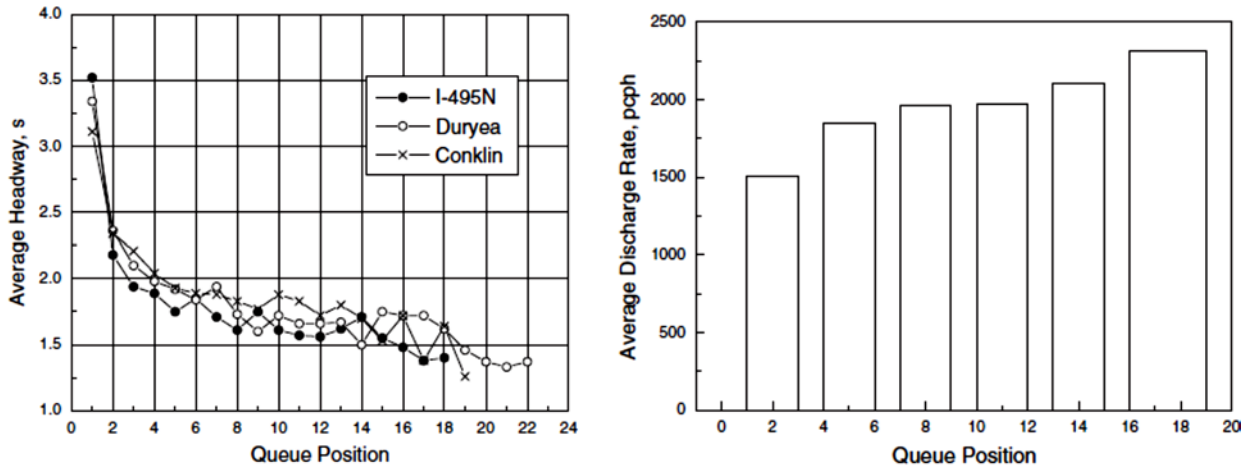
Figure 14: Variation in Headway Distribution by Measurement Location [22]

Nevertheless, other studies, based on observed field data, showed that queue discharge may not display stable saturation flow rates. Instead, by progressing the queue discharge the headway have a tendency to compress continuously. Consequently, the saturation flow rate keeps rising and does not reach a constant rate. Some of these studies argued that the HCM procedures may be based on some built-in errors, as discussed hereinafter.

One of these studies was conducted by Lin and Thomas [25], which disagrees with the HCM concept that the queue discharge headways become constant after the fourth vehicle. Contrarily, they argue that the queue discharge headways continue to get smaller even after the 15th stopped vehicle in the queue discharges [25]. As a result, the queue discharge rates keep rising with the progress of the queue. The study concluded its point of view based on field observations from three straight-through traffic lanes at intersections in New York City. The relationship between the average queue discharge headway and the queue position from these observations is shown in Figure 15 (a). According to the figure, the headway does not appear to converge to some stable headway after the fourth vehicle, instead the headways are gradually compressed as queuing vehicles discharge in succession [25].

Moreover, the relation of average discharge rate with queue position is determined for each intersection, as shown in Figure 15 (b). As the sample size for some queue positions was small, the average of three to four queue positions were grouped and the average discharge rate was determined to attain better perspective [25]. As demonstrated from the figure the average queue discharge rate

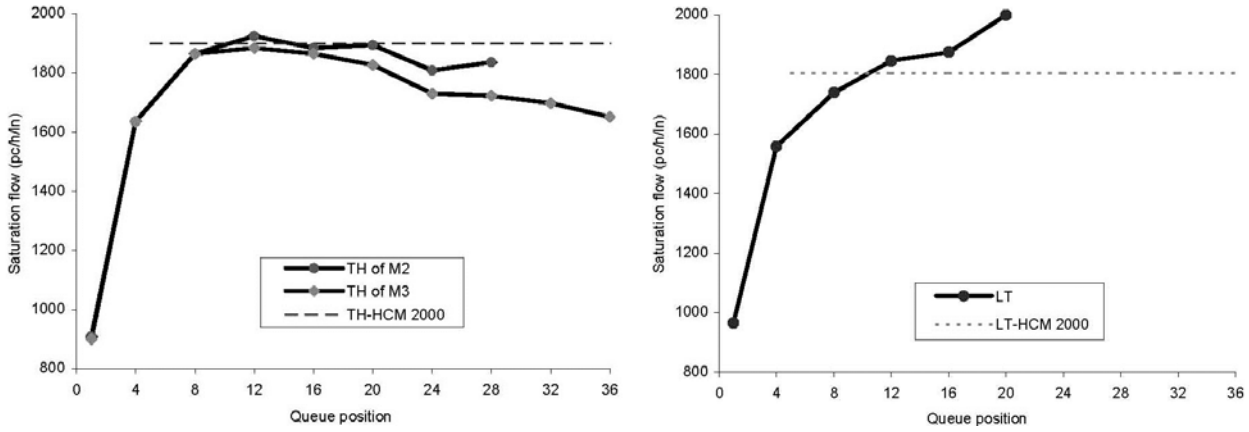
increases from one group to the other. The differences between any two successive groups were found significant except between the third and fourth group there was no change in the discharge rate. Throughout both figures, the queue discharge characteristics at all the three intersections reveal a gradual trend of headway compression with the continuing of the queue discharge which does not match the HCM methodology. Lin and Thomas suggested that the methodology in the HCM may include built-in errors for the congestion analysis at many signalized intersections around the United States [25]. Reducing the errors in the estimation of capacity is very important as these errors can lead to larger errors in the estimation of vehicle delays and intersection level of service classification. Contrarily, the accurate estimation of vehicle delay is very important to correctly classify the level of service and to make good design and operating decisions [25].



(a) (b)

Figure 15: Queue Discharge Headways (a) and Rates (b) at Three Intersections in NY [25]

Another study by Li and Prevedouros [26] analyzed headway distributions for through and left-turn movements at one intersection in Honolulu, HI. Similar to Lin and Thomas [25], the study concluded that the headways and the corresponding saturation flows do not reach constant values, as presented in Figure 16. Another interesting finding by the paper is that the HCM overestimated the saturation flow for through movements and underestimated it for left-turn movements. Also, longer green times with longer queues appeared to be less productive with decreased saturation flows [27].



(a) (b)

Figure 16: Saturation Flows for Through (a) and Left-turn (b) Movements in HI [27]

In contrast to what was concluded from Lin and Thomas [25], and Li and Prevedouros [26], Cohen [27] in his study, queue discharge at signalized intersections was analyzed using simulation analysis with application of the Pitt car-following model. A number of different scenarios were executed in order to examine the impact of different factors. These factors include free-flow speed, car following parameters, vehicle length, traffic stream composition, and lane changing. It was concluded that all of these factors have significant effects on the discharge headways. For example, the increase of the free-flow speed yields a decrease in the discharge headways, as shown in Figure 17. In general, the study concluded that the discharge headway distribution is almost flat beyond the fifth vehicle in the queue [27]. This is consistent to the HCM assumption mentioned earlier, but contradicts Lin and Thomas [25], and Li and Prevedouros [26] findings.

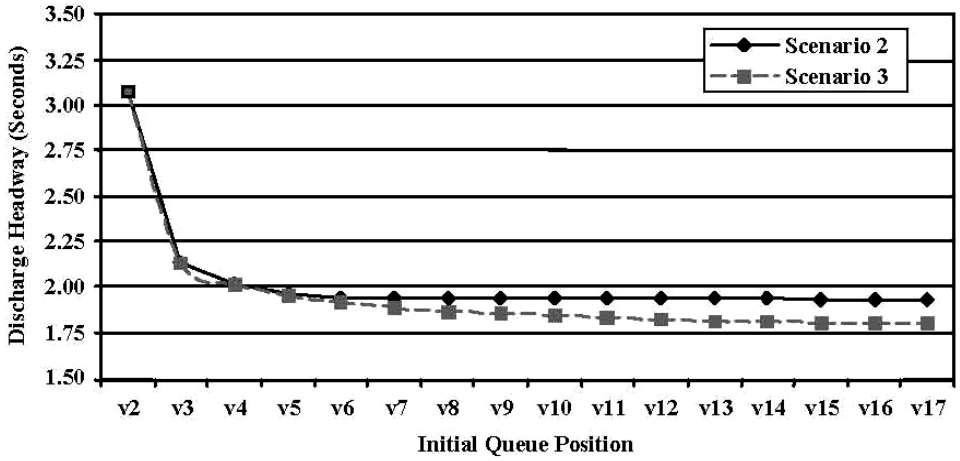


Figure 17: Headways for Low (Scenario 2) and High (Scenario 3) Free Flow Speeds [27]

2.4 SUMMARY AND CONCLUSIONS

This chapter synthesized the available literature related to the issue of capacity drop and flow reduction after the onset of congestion. The literature are typically divided into those studies related to capacity drop on uninterrupted flow facilities (i.e. freeways), and those related to drop in saturation flow rate on interrupted flow facilities (e.g. signalized intersections).

The capacity drop at the onset of congestion at freeway bottlenecks is well documented and has been empirically demonstrated. The losses reported in the literature range from 5 to 26% depending on the literature source. In the case of signalized intersections there is agreement that there are losses in saturation flow rates initially, however there appears to be disagreement on whether these losses are spatially and temporally confined. A number of issues need further investigation. First, what is the reason for the observed losses? One reason presented relates to vehicle acceleration constraints. However, it is not clear why the losses associated with a backward recovery shockwave are spatially and temporally confined. Consequently, the research presented in this thesis attempts to address the following issues:

1. Can the field observed capacity drop be replicated in an agent-based modeling approach that accounts for vehicle acceleration constraints?
2. How sensitive is the capacity drop to the level of vehicle acceleration?
3. Are the reductions in flow discharge rates from a backward recovery wave spatially and temporally confined? If so why are they different from stationary waves?
4. What factors affect these losses and how sensitive are the losses to these factors?

CHAPTER 3 : SIMULATION STUDY OF FREEWAY BOTTLENECK FLOW REDUCTIONS AT THE ONSET OF CONGESTION

Edward Chamberlayne[1], Hesham Rakha[2], Maha El-Metwally[3], and Douglas Bish[4]
(accepted for presentation at the 2011 TRB 90th Annual Meeting)

3.1 ABSTRACT

Empirical studies have demonstrated that the discharge flow rate at a bottleneck is reduced following the onset of congestion. These flow reductions are typically measured by comparing the queue discharge to the maximum pre-queue flow rate. The paper demonstrates, through the use of the INTEGRATION microscopic traffic simulation software, that these empirically observed reductions in flow can be simulated without enforcing a discontinuity in the fundamental diagram. Instead these reductions can be captured by constraining vehicle acceleration levels, which in turn introduces a discontinuity in the data. The study demonstrates that INTEGRATION produces flow reductions in the same level of magnitude as empirically observed. Specifically, reductions in the range of 10 to 15 percent for on-ramp bottlenecks and reductions in the range of 2 to 5 percent for lane reduction bottlenecks are observed. Furthermore, as was empirically observed, these reductions are higher for on-ramp bottlenecks compared to lane reduction bottlenecks. The study demonstrates that the flow reductions are very sensitive to the level of acceleration that drivers are willing to exert and the percentage of heavy vehicles in the traffic stream.

3.2 INTRODUCTION

The fact that traffic congestion upstream of a freeway bottleneck reduces the discharge flow from the bottleneck has been well documented in several empirical studies. The problem in many of the empirical studies, however, is that it is very difficult to determine what exactly caused the congestion and, in turn, what caused the flow reductions downstream. Additionally, when simulating traffic flows, it is important to understand the contribution of various factors on these flow reductions.

In order to quantify the flow reduction, the downstream flow during congestion has to be compared to a baseline flow condition prior to the formation of congestion. This is a difficult task for

[1] Ph.D. Candidate, Grado Department of Industrial & Systems Engineering, Virginia Tech
[2] Professor, Charles E. Via, Jr. Department of Civil and Environmental Engineering, Virginia Tech
[3] M.S. Student, Charles E. Via, Jr. Department of Civil and Environmental Engineering, Virginia Tech
[4] Assistant Professor, Grado Department of Industrial & Systems Engineering, Virginia Tech

empirical studies and can be highly dependent on time-of-day, year, weather conditions, and other factors.

Many of the papers that discuss these reductions refer to the "capacity drop" phenomenon which refers to the original two capacity phenomenon introduced by Edie [12]. This phenomenon requires a discontinuity within the fundamental diagram. This paper will use the term "flow reduction" since the flow is reduced irrespective whether it is a result of a discontinuity in the fundamental diagram or discharging from a different location along the fundamental diagram.

This paper will attempt to explore the sources of flow reductions through the use of a microscopic traffic simulator, INTEGRATION. By using the simulation software, the composition of the traffic stream and the behavior of the traffic in terms of vehicle dynamics can be controlled. Through this level of control, the flow reductions can be quantified more accurately and attributed to the various sources of congestion. The baseline flow conditions can equally be controlled and various situations can be modeled within a simulation environment.

The paper will briefly review the current applicable literature and empirical studies, then present two common freeway bottleneck examples, discuss the results, and present some conclusions related to the results of the simulations.

3.3 BACKGROUND

The concept of two capacities or the "capacity drop" was first introduced by Edie in 1961 [12] where congested and uncongested flows in the Lincoln Tunnel in New York were analyzed. This empirical study gave the first support to the idea of flow reductions downstream of a congested bottleneck. In Hall and Agyemang-Duah 1991 [7], they concluded that once a queue forms upstream of a bottleneck, a capacity drop appears as a consequence of the way drivers accelerate away from the queue. Their conclusion was based on data collected from the Queen Elizabeth Way outside of Toronto and the results showed a capacity drop of approximately 5 to 6 percent [7]. In addition to the idea of capacity drop, the paper argued that the capacity drop should be measured within the bottleneck – not upstream or downstream of the bottleneck [7]; otherwise only a reduction in flow would be measured. The authors demonstrated that measuring at this location yields a one branch occupancy-flow relationship. This branch consists of pre-queue flows and queue discharge flows. The vertical difference between these two curves presents the capacity drop. In this study, a truck percentage of 6 percent was reported in the traffic stream [7].

Cassidy and Bertini 1999 [17] studied the capacity drop phenomenon along the QEW as was done in the previous study [7]. A capacity drop of 8 to 9 percent was found along the QEW, as compared to the 5 to 6 percent drop found in [7], by visually comparing transformed cumulative curves and cumulative occupancy against time created by data from downstream loop detectors [17].

One more argument about the issue of capacity drop was raised by Banks in 1991 [18]. Even though the study concluded that bottleneck capacities decrease when flow breaks down due to congestion, it stated that this applies to the capacities of individual lanes [18]. Four different bottlenecks in four sites in San Diego were investigated. It was shown that capacity decreased after breakdown by about 10 percent in the left lane. However, the capacity, when averaged across all lanes, decreased by only 3 percent.

Chung et al. in 2007 inferred that an increase in the freeway density is always followed by a drop in capacity [1]. In their study, three different bottlenecks on different freeway segments in San Diego and San Francisco were investigated to examine the relationship between vehicular density and the drop in capacity. The three bottlenecks investigated were an on-ramp merge, a reduction in the number of travel lanes, and a horizontal curve on the freeway. The study concluded that a flow reduction of 10 percent resulted from the formation of a queue in the freeway shoulder lane upstream of the merge. A critical density of about 208 vehicles/km was assumed as a threshold corresponding to the capacity drop [1]. A capacity drop of 5 percent was reported for the lane reduction, which also occurred when the density reached a critical density of approximately 90 vehicles/km [1]. Nevertheless, the conclusion that a specific value for critical density causes the capacity drop requires further investigation. Generalizing a fixed critical density value to be applied at any freeway could not be accepted unconditionally, as the capacity drop can result from factors other than increased density. Contrarily, increased traffic density could occur without corresponding drop in capacity. For instance, a sudden lane changing decision from one driver can block a lane and increase its density, while the traffic volume is not high or near capacity. This discussion implies that not only high traffic density can trigger capacity drop, but it should be accompanied by high flow rates. In spite of that, no deterministic value for density or flow can accompany the reduction in capacity. In more than one study [13, 14], Kerner investigated the flow breakdown phenomenon and agreed that the breakdown is a probabilistic phenomenon, and not deterministic as concluded by Chung et al. [1].

Kerner refers to an earlier study by Persaud et al. [3], where it was stated that the occurrence of a breakdown has a probabilistic nature. To explore the probability of breakdown, traffic data were collected at three different sites in Toronto, Canada using surveillance cameras and automatic collection

of speed and flow data from loop detectors. A capacity drop was found ranging from 11 percent to 17 percent at two of the sites, while at the third site the capacity drop was 26 percent [3]. Furthermore, the probability of breakdown at various traffic flow levels was investigated [3]. In general for the three sites, the probability of breakdown increased by increasing the flow rates. It was found that by maintaining the pre-queue flows at the same level as those that occur after a queue forms, the probability of breakdown was almost negligible. By increasing these flows by 20 percent above the queue discharge flow, the probability of breakdown was only 10 percent. For more increased flows, the probability of breakdown increased dramatically. Hence, it was suggested not to operate at pre-queue rates exceeding the queue discharge rate by more than 20 percent [3].

Given that one explanation for the reduction in discharge flow after the onset of congestion relates to driver acceleration levels, a review of the literature in this area was also conducted. A study by Snare 2002, found that drivers normally accelerate at 60 percent of the maximum acceleration level of their vehicle; “Driver factors were observed that fit a normal distribution with a mean of 0.6 and a standard deviation of 0.08” [20]. This confirmed a previous guideline published in 1954 by AASHO [21]. This conclusion is important since the flow reductions are influenced by how vehicles accelerate downstream of the bottleneck.

3.4 INTEGRATION SOFTWARE OVERVIEW

This section provides a brief overview of the INTEGRATION software given that it was utilized to conduct the study. The INTEGRATION software is a microscopic traffic assignment and simulation software that was developed over the past decade [28, 29, 30, 31]. It was conceived as an integrated simulation and traffic assignment model and performs traffic simulations by tracking the movement of individual vehicles every 1/10th of a second. This allows detailed analyses of lane-changing movements and shock wave propagations. It also permits considerable flexibility in representing spatial and temporal variations in traffic conditions. In addition to estimating stops and delays [32, 33, 34], the model can also estimate the fuel consumed by individual vehicles, as well as the emissions [35, 36, 37, 38]. Finally, the model also estimates the expected number of vehicle crashes using a time-based crash prediction model [39].

The INTEGRATION model updates vehicle speeds every deci-second based on a user-specified steady-state speed-spacing relationship and the speed differential between the subject vehicle and the vehicle immediately ahead of it. In order to ensure realistic vehicle accelerations, the model uses a vehicle dynamics model that estimates the maximum vehicle acceleration level. The INTEGRATION car-

following model falls within the psycho-physical model formulations because it considers a driver within different regimes. Specifically, the vehicle will collide with a vehicle if its spacing from the lead vehicle is less than a safe distance that increases as the speed difference between the lead and following vehicle increases (negative speed differential). The driver typically attempts to converge to the steady-state behavior by either decelerating (collision avoidance) or accelerating. The driver is in free-flow mode once the spacing between the following and lead vehicle exceeds a threshold.

The model computes the vehicle speed as the minimum of two speeds, namely: the maximum vehicle speed based on vehicle dynamics and the desired speed based on the Van Aerde steady-state car-following model formulation as

$$u_n(t+\Delta t)=\min\left\{\begin{array}{l} u_n(t)+3.6\cdot\dfrac{F_n(t)-R_n(t)}{m}\Delta t, \\ \dfrac{-c_1'+c_3u_f+\tilde{s}_n(t)-\sqrt{\left[c_1'-c_3u_f-\tilde{s}_n(t)\right]^2-4c_3\left[\tilde{s}_n(t)u_f-c_1'u_f-c_2\right]}}{2c_3} \end{array}\right\} \quad (1)$$

3.4.1 Steady-State Modeling

The Van Aerde nonlinear single-regime functional form that was proposed by Van Aerde [40] and Van Aerde and Rakha [41], is formulated as

$$s_n(t)=c_1+c_3u_n(t+\Delta t)+\frac{c_2}{u_f-u_n(t+\Delta t)}, \quad (2)$$

where c_1, c_2, and c_3 are model constants. Demarchi [42] demonstrated that by considering three boundary conditions the model constants can be computed as

$$c_1=\frac{u_f}{k_ju_c^2}\left(2u_c-u_f\right);\quad c_2=\frac{u_f}{k_ju_c^2}\left(u_f-u_c\right)^2;\quad c_3=\left(\frac{1}{q_c}-\frac{u_f}{k_ju_c^2}\right). \quad (3)$$

As was demonstrated by Rakha and Crowther [43] this functional form amalgamates the Greenshields and Pipes car-following models.

Ignoring differences in vehicle behavior within a traffic stream and considering the relationship between traffic stream density and traffic spacing, the speed-density relationship can be derived as

$$k=\frac{1000}{c_1+\dfrac{c_2}{u_f-u}+c_3u}, \quad (4)$$

Of interest is the fact that Equation (4) reverts to Greenshields' linear model, when the speed-at-capacity and density-at-capacity are both set equal to half the free-flow speed and jam density,

respectively (i.e. u_c=u_f/2 and k_c=k_j/2). Alternatively, setting u_c=u_f results in the linear Pipes model given that

$$c_1 = \frac{1}{k_j} = s_j; \quad c_2 = 0; \quad c_3 = \frac{1}{q_c} - \frac{1}{k_j u_f}.$$

Rakha [44] demonstrated that the wave speed at jam density (denoted as w_j) can be computed as

$$w_j = -s_j \left. \frac{1}{\frac{ds}{du}} \right|_{u=0} = -\frac{s_j}{c_3 + \frac{c_2}{u_f^2}} = -\frac{u_f}{k_j\left(c_3 u_f^2 + c_2\right)} = -\left[\left(\frac{k_j}{q_c} - \frac{u_f}{u_c^2}\right) + \frac{\left(u_f - u_c\right)^2}{u_f u_c^2}\right]^{-1}. \quad (5)$$

3.4.2 Vehicle Deceleration Behavior

The literature [45] indicates that the maximum braking force acting on each axle can be computed as the coefficient of roadway adhesion multiplied by the vehicle weight normal to the roadway surface. Because true optimal brake force proportioning is seldom achieved in standard non-antilock braking systems, a braking efficiency term is also used in computing the maximum braking force as

$$d_{max} = \eta_b \mu g. \quad (6)$$

Here η_b is the braking efficiency, μ is the coefficient of roadway adhesion also known as the coefficient of friction, and g is the gravitational acceleration (9.8066 m/s^2). In the case of antilock braking systems the braking efficiency approaches 100 percent. Noteworthy is the fact that Equation (6) demonstrates that the maximum vehicle deceleration varies as a function of the roadway conditions as reflected by the coefficient of road friction.

The INTEGRATION model ensures that a vehicle maintains an additional distance to the lead vehicle to allow the following vehicle to decelerate to the speed of the lead vehicle in the event that the following vehicle is closing on the lead vehicle.

3.4.3 Vehicle Acceleration Modeling

Vehicle acceleration is governed by vehicle dynamics. Vehicle dynamics models compute the maximum vehicle acceleration levels from the resultant force acting on a vehicle, as

$$a = \frac{F - R}{m} \quad (7)$$

where a is the vehicle acceleration (m/s^2), F is the vehicle tractive force (N), R is the total resistance force (N), and m is the vehicle mass (kg).

The vehicle tractive effort is computed as

$$F_T = 3600\, f_p \beta \eta \frac{P}{u}. \tag{8}$$

Here F_T is the engine tractive force (N), f_p is the proportion of the maximum acceleration that the driver is willing to employ (field studies have shown that it is typically 0.60), β is a gear reduction factor that will be described later (unitless), η is the driveline efficiency (unitless), P is the vehicle power (kW), and u is the vehicle speed (km/h).

Given that the tractive effort tends to infinity as the vehicle speed tends to zero, the tractive force cannot exceed the maximum force that can be sustained between the vehicle's tractive axle tires and the roadway surface, which is computed as

$$F_{max} = m_{ta}\, g \mu. \tag{9}$$

Here m_{ta} is the mass of the vehicle on the tractive axle (kg), g is the gravitational acceleration (9.8066 m/s^2), and μ is the coefficient of road adhesion or the coefficient of friction (unitless).

Typical axle mass distributions for different truck types were presented in an earlier publication and thus are not discussed further [46]. The tractive force is then computed as the minimum of the two forces as

$$F = \min(F_T, F_{max}). \tag{10}$$

Rakha and Lucic [47] introduced the β factor into Equation (8), in order to account for the gear shift impacts at low traveling speeds when trucks are accelerating. Specifically, the factor is a linear function of vehicle speed with an intercept of $1/u_0$ and a maximum value of 1.0 at u_0 (optimum speed or the speed at which the vehicle attains its full power) as

$$\beta = \frac{1}{u_0}\left[1 + \min\left(u, u_0\right)\left(1 - \frac{1}{u_0}\right)\right]. \tag{11}$$

The optimum speed was found to vary as a function of the weight-to-power ratio (for weight-to-power ratios (w) ranging from 30 to 170 kg/kW) as

$$u_0 = 1164 w^{-0.75}. \tag{12}$$

Here w is the weight-to-power ration in kg/kW. Rakha and Snare [48] demonstrated that the gear shift parameter β is not required for the modeling of light-duty vehicle acceleration behavior (weight-to-power is less than 30 kg/kW).

Three resistance forces are considered in the model, namely the aerodynamic, rolling, and grade resistance forces [45, 46]. The first resistance force is the aerodynamic resistance that varies as a function of the square of the air speed. Although a precise description of the various forces would involve the use of vectors, for most transportation applications scalar equations suffice if the forces are considered to only apply in the roadway longitudinal direction. For the motion of a vehicle in still air, the air speed equals the vehicles speed as

$$R_a = \frac{\rho}{2 \times 3.6^2} C_d C_h A u^2 = c_1 C_d C_h A u^2 , \quad (13)$$

where ρ is the density of air at sea level and a temperature of 15ºC (59ºF) (equal to 1.2256 kg/m^3), C_d is the drag coefficient (unitless), C_h is a correction factor for altitude (unitless), and A is the vehicle frontal area (m^2). Given that the air density varies as a function of altitude, the C_h factor can be computed as

$$C_h = 1 - 8.5 \times 10^{-5} H . \quad (14)$$

Typical values of vehicle frontal areas for different vehicle types and typical drag coefficients are provided in the literature [46].

The second resistance force is the rolling resistance, which is a linear function of the vehicle speed and mass, as

$$R_r = C_r (c_2 u + c_3) \frac{mg}{1000} . \quad (15)$$

Typical values for the rolling coefficients (C_r, c_1, and c_2), as a function of the road surface type, condition, and vehicle tires, are provided in the literature [46]. Generally, radial tires provide a resistance that is 25 percent less than that for bias ply tires.

The third and final resistance force is the grade resistance, which accounts for the proportion of the vehicle weight that resists the movement as a function of the roadway grade (i) as

$$R_g = mgi . \quad (16)$$

Having computed the various resistance forces, the total resistance force is computed as

$$R = R_a + R_r + R_g . \quad (17)$$

3.5 EXPERIMENTAL DESIGN

In order to study and characterize the empirically observed flow reduction phenomenon, two simple examples are modeled. Before describing the specific examples, the following section describes the general assumptions made and the scenarios considered in each example.

3.5.1 Assumptions

Several assumptions were made in order to set criteria to measure downstream flow reductions during the presence of congestion. In order to replicate the results from previous empirical studies, a 30-second sampling interval was used, as used in [7]. The study compiled average flow and occupancy measurements at loop detectors at both the stations upstream and downstream of the bottleneck.

In determining the onset of congestion, occupancy measurements at the upstream station were monitored. Congestion was considered to have formed once the occupancy exceeded a specific threshold for a 3-minute time period. A 3-minute time period was considered in order to be consistent with earlier empirical studies [7]. Equation (18) gives the relationship between the critical density (k_c) and average vehicle length ($\overline{L_v}$), loop detector length (L_D), and detector occupancy (O_D).

$$k_c = \frac{1000}{\overline{L_v} + L_D} O_D \tag{18}$$

$$k_c = \frac{q_c}{u_c} = \frac{q_c}{u_f} \tag{19}$$

$$O_D = \frac{q_c \left(\overline{L_v} + L_D \right)}{1000\, u_f} \tag{20}$$

Equation (19) calculates k_c by dividing the saturation flow rate (or capacity), q_c, by the speed-at-capacity. The speed-at-capacity, u_c, was set equal to the free-flow speed, u_f, to model the commonly known triangular flow-density relationship (fundamental diagram) or the Pipes car-following model. By substituting Equation (19) into Equation (18) and solving for the detector occupancy, a threshold for the start of congestion can be solved for. Using the average vehicle length of 4 m, a detection zone length of 5 m, a capacity of 2400 veh/h/lane, and free-flow speed of 108 km/h, the threshold for detector occupancy, O_D, corresponds to 20 percent. The onset of congestion corresponded to the instant at which the upstream detector occupancy exceeded 20 percent for at least a 3-minute period. The corresponding downstream flows were temporally offset by the travel time from the upstream to the downstream detector at free-flow speed. Flow reductions were made relative to the base downstream flow, as will be described later.

Two vehicle types were used in the simulation runs; a light duty composite vehicle to represent passenger cars and a medium truck to represent heavy-duty vehicles. The light duty vehicle has a mass of 1326 kg, is 4.8m in length, and 109kW of power. The medium truck has a mass of 31,751 kg, is 16m in length, and 261 kW of power.

Lastly, a vehicle speed coefficient of variation of 5 percent was coded within the INTEGRATION software to capture differences in driver car-following behavior.

3.5.2 Scenarios

For the two freeway bottleneck examples, four scenarios were considered.

- Scenario 1 – all vehicles are light duty cars and accelerate at 100 percent of their maximum capability. This corresponds to a vehicle acceleration at maximum throttle.
- Scenario 2 – all vehicles are light duty cars and accelerate at the typical acceleration level of 60 percent of their maximum capability.
- Scenario 3 – traffic stream includes 6 percent heavy vehicles and all vehicles accelerate at the typical acceleration of 60 percent of the maximum level.
- Scenario 4 – traffic stream includes 15 percent heavy vehicles and all vehicles accelerate at the typical acceleration of 60 percent their maximum level.

The change from 100 percent to 60 percent of the maximum acceleration capability was inspired by [20] and previous research cited therein. The 6 percent heavy vehicle composition was taken from the empirical data used in [7, 17].

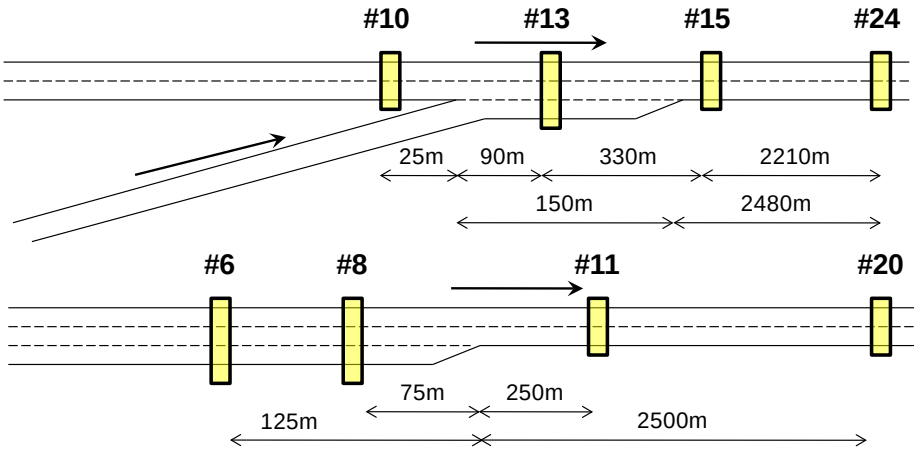


Figure 18: (a) Freeway On-ramp Bottleneck (top); (b) Freeway Lane Reduction Bottleneck (bottom)

3.5.3 On-ramp Bottleneck Example

The first bottleneck example that was modeled is a freeway merge section, shown in Figure 18(a). This example considers a 2-lane freeway segment with a 1-lane on-ramp merging onto the freeway. Loop detectors were installed every 100 meters, however data from station 10 and 24 are only reported to capture conditions both upstream and downstream of the bottleneck.

The following macroscopic parameters were used to simulate this example. For the freeway, q_c = 2400 veh/hour/lane and u_f = 108 km/hr. For the on-ramp, q_c = 1800 veh/hour/lane and u_f = 72 km/hr. A jam density (k_j) of 150 veh/km/lane was used for both the freeway and on-ramp. The speed-at-capacity, u_c, is assumed to equal u_f (Pipes model). It should be noted that the capacity that is coded is the base roadway capacity without considering the impacts of lane-changing, trucks, and variability in driver behavior on the the capacity. The actual capacity is typically lower than the base capacity.

For each of the four scenarios, the freeway demand was fixed at 3600 vph while the on-ramp demand was increased at increments of 100 vph from an initial demand of 100 vph to 1200 vph. The maximum on-ramp flow that did not result in the formation of congestion (occupancy of 20 percent for 3 minutes or longer) at the upstream detector station (detector station 10) was used to establish the base pre-queue maximum flow. The average flow for the pre-queue scenario at the downstream detector was then used as the base reference flow for the computation of discharge flow reductions. This approach is consistent with ramp metering logic where the on-ramp flow is controlled in order to ensure that no breakdown occurs on the freeway mainline (i.e. no congestion forms upstream of the on-ramp). The baseline flow corresponded to a loading of 3600 (freeway) /500 (on-ramp) vph for Scenario 1, 3600/300 vph for Scenario 2, and 3600/100 vph for Scenarios 3 and 4.

3.5.4 Lane Reduction Bottleneck Example

The second example is a freeway lane reduction where 3-lanes reduce to 2-lanes, as shown in Figure 18(b). As in merge bottleneck example, loop detectors were located every 100 m but only data from stations 6 and 20 are reported to capture conditions both upstream and downstream of the bottleneck.

The following macroscopic parameters were used to simulate this example:base q_c = 2400 veh/hour/lane, u_f = 108 km/hr, and k_j = 150 veh/km/lane.

As was the case with the on-ramp bottleneck, several different demand loading profiles were tested in each of the four scenarios. The freeway demands used are as follows: 4500, 4800, 5400, 6000,

6600, and 7200 vph. Note that the base capacity of the 2-lane section is 4800 vph. Two separate baseline flow conditions were used in this example, as follows:

- **Baseline 1:** The average downstream flow at Station 20 was used as a baseline when the network was loaded with enough demand that would NOT result in congestion upstream of the bottleneck at Station 6. The baseline flow conditions were established by loading 4500 vph of demand for Scenario 1, 4200 vph for Scenario 2, 3600 vph for Scenario 3, and 2900 vph for Scenario 4. As is expected these capacities decrease as vehicles accelerate less aggressively and as trucks are introduced into the traffic stream.
- **Baseline 2:** The average downstream flow at Station 20 when the entire section is composed of two lanes. This base flow eliminates any lane changing behavior effects on the baseline flow. This was intended to simulate traffic behavior along straight freeway sections when the traffic consolidates into 2-lanes well before the lane reduction from 3-lanes to 2-lanes occurs. Because of the randomness in the simulations, the throughput for these cases could exceed the base capacity of 4800 vph. In the case when the througput exceeded the base capacity of the bottleneck, the bottleneck capacity was used as the baseline flow.

3.6 STUDY FINDINGS

The next section summarizes the results of the simulations for each of the two bottleneck examples. The simulation results reveal reductions in discharge flow rates after the onset of congestion for each of the two example bottleneck examples. The following sections first discuss the validity of the INTEGRATION software, then analyze and discuss the results for each bottleneck/scenario combination.

3.6.1 Validation of the INTEGRATION Software

In order to validate the INTEGRATION software and its ability to simulate the capacity drop phenomenon, the results of scenario 3 presented in Table 1 were used. First, in scenario 3, a truck percentage of 6% is simulated in the traffic stream in order to replicate what was done in [7]. In addition, a typical acceleration level of 60% is used to capture typical levels of vehicle acceleration as was presented earlier in [20]. As shown in Table 1, nine different demand loading cases are used in scenario 3. Congestion forms once the total demand is 3800 vph, which corresponds to a flow reduction of 8%. By varying the demand loading case, the discharge flow reductions are in the range of 8 to 19%. These results are consistent with the results found previously in the literature [1, 7, 17, 19].

3.6.2 On-ramp Bottleneck Results

The four scenarios indicate that the flow reduction created by the formation of congestion upstream of an on-ramp bottleneck is sensitive to several different contributors, as summarized in Table 1.

Table 1: Flow Reductions for the On-ramp Bottleneck

	Demand Loading Case								
	1	2	3	4	5	6	7	8	9
Scenario 1: Maximum Acceleration with Cars Only									
Baseline Flow (vph)	*N/A*	*N/A*	*N/A*	*N/A*	4088	4088	4088	4088	4088
Avg. Queue Discharge Flow (vph)	*N/A*	*N/A*	*N/A*	*N/A*	4088	4114	4084	4034	4065
Flow Reduction (%)	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	-0.7 %	0.1 %	1.3 %	0.5 %
Freeway Demand (vph)	3600	3600	3600	3600	3600	3600	3600	3600	3600
On-ramp Demand (vph)	100	200	300	400	500	600	800	1000	1200
On-ramp Demand (%)	3 %	5 %	8 %	10 %	12 %	14 %	18 %	22 %	25 %
Scenario 2: Typical Acceleration with Cars Only									
Baseline Flow (vph)	N/A	N/A	3851	3851	3851	3851	3851	3851	3851
Avg. Queue Discharge Flow (vph)	N/A	N/A	3851	3789	3704	3643	3564	3562	3550
Flow Reduction (%)	N/A	N/A	N/A	1.6 %	3.8 %	5.4 %	7.5 %	7.5 %	7.8 %
Freeway Demand (vph)	3600	3600	3600	3600	3600	3600	3600	3600	3600
On-ramp Demand (vph)	100	200	300	400	500	600	800	1000	1200
On-ramp Demand (%)	3 %	5 %	8 %	10 %	12 %	14 %	18 %	22 %	25 %
Scenario 3: Typical Acceleration with 6 % Trucks									
Baseline Flow (vph)	3615	3615	3615	3615	3615	3615	3615	3615	3615
Avg. Queue Discharge Flow (vph)	3615	3326	3128	3089	3071	2938	2911	2934	2982
Flow Reduction (%)	N/A	8.0 %	13.5 %	14.5 %	15.0 %	18.7 %	19.5 %	18.8 %	17.5 %
Freeway Demand (vph)	3600	3600	3600	3600	3600	3600	3600	3600	3600
On-ramp Demand (vph)	100	200	300	400	500	600	800	1000	1200
On-ramp Demand (%)	3 %	5 %	8 %	10 %	12 %	14 %	18 %	22 %	25 %
Scenario 4: Typical Acceleration with 15 % Trucks									
Baseline Flow (vph)	2912	2912	2912	2912	2912	2912	2912	2912	2912
Avg. Queue Discharge Flow (vph)	2912	2738	2595	2516	2434	2464	2441	2450	2377
Flow Reduction (%)	N/A	6.0 %	10.9 %	13.6 %	16.4 %	15.4 %	16.2 %	15.9 %	18.4 %
Freeway Demand (vph)	3600	3600	3600	3600	3600	3600	3600	3600	3600
On-ramp Demand (vph)	100	200	300	400	500	600	800	1000	1200
On-ramp Demand (%)	3 %	5 %	8 %	10 %	12 %	14 %	18 %	22 %	25 %

The flow reductions are sensitive to the magnitude of total demand loaded and the proportion of the total demand entering from the on-ramp. This is not surprising since as the on-ramp demand increases, the total demand increases, which causes the amount of congestion to increase upstream of the merge. In Scenario 1, congestion does not form until the total demand is 4200 vph and the on-ramp demand comprises 14 percent of the total demand. However, congestion forms at 10, 5, and 5 percent of the total demand for Scenario 2, 3 and 4, respectively.

Another factor in the creation of flow reductions is the acceleration level used by the drivers. The flow reductions in Scenario 1 are rather small, less than 2 percent. However, once the acceleration level of the traffic stream is decreased from 100 to 60 percent of the maximum rate (as demonstrated in [20] as a more typical value), the flow reductions increase. The flow reductions in Scenario 2 are in the range of 1 to 8 percent which are substantially larger than those reductions given in Scenario 1. Additionally, as mentioned in the previous paragraph, congestion builds at an on-ramp demand of 10 percent of the total demand instead of the 14 percent demand in Scenario 1.

The remaining significant factor in the creation of flow reductions is the percentage of heavy vehicles within the traffic stream. Scenarios 3 and 4 give the flow reductions when heavy vehicle percentages of 6 and 15 percent are introduced, respectively. In Scenario 3, the flow reductions begin at an on-ramp demand of just 5 percent, as compared to 10 percent in Scenario 2. The discharge flow reductions are in a range of 8 to 19 percent, which is approximately a 2-fold increase from the flow reductions in Scenario 2. The introduction of 6 percent heavy vehicles creates a decrease in the average downstream flow due to their decreased acceleration capability and the fact that they delay passenger cars following behind them and/or passing them (the moving bottleneck effect). Interestingly, the flow reductions in Scenario 4 are slightly less than those from Scenario 3. However, the average flows downstream from Scenario 4 are less than those from Scenario 3. The reductions are in the range of 6 to 19 percent as compared to the 8 to 19 percent range from Scenario 3. These results are a little counter-intuitive since it could be argued that the flow reductions should be higher when the percentage of heavy vehicles is higher. However, it appears that the effect of trucks stabilizes when the heavy vehicle percentage is increased from 6 percent. This behavior is consistent with findings in the literature [49]. However, the flow reductions from Scenario 4 are significantly greater than those from Scenario 2 (6 to 19 percent reductions versus 1 to 8 percent).

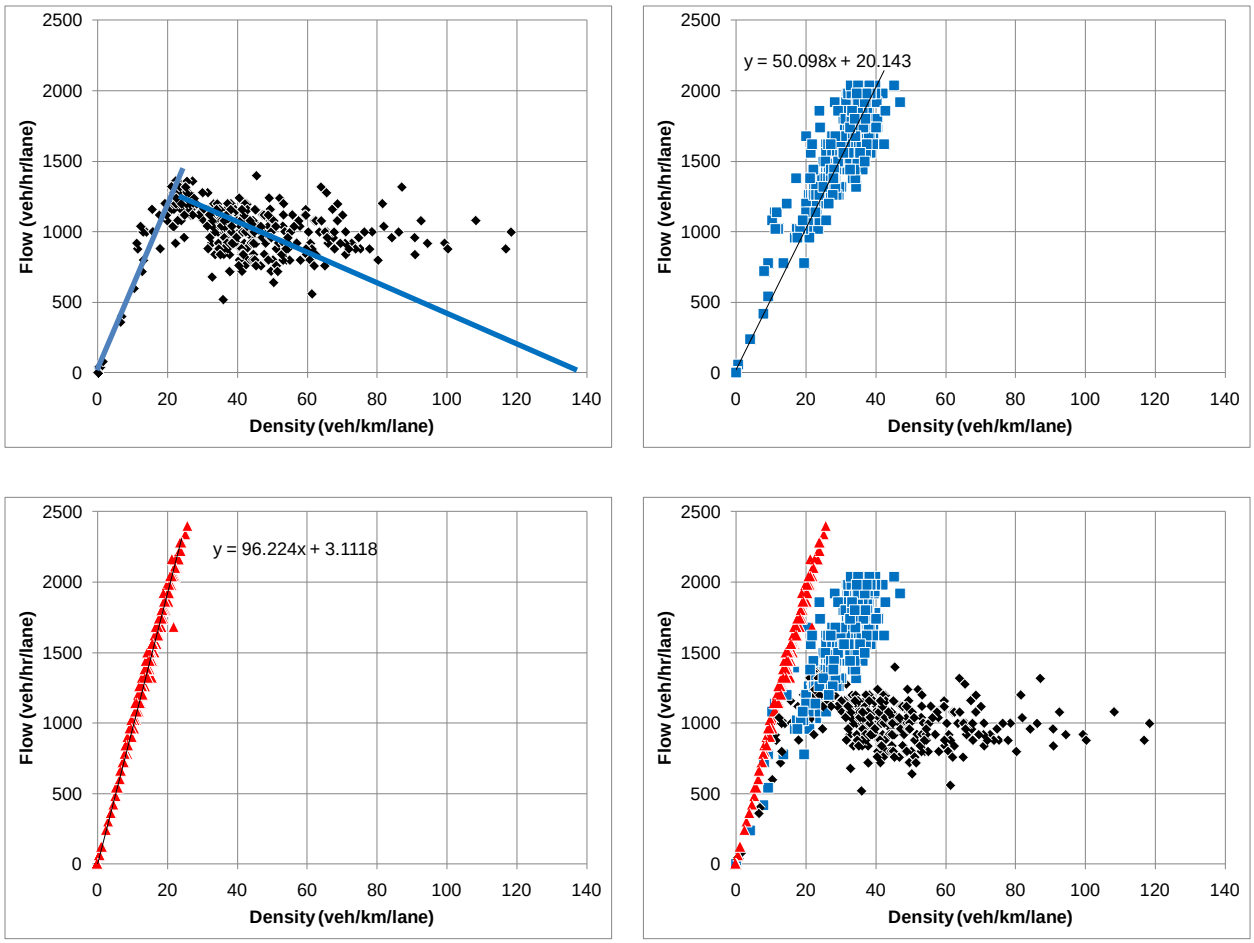


Figure 19: Flow-Density Plots for the On-ramp Bottleneck for Scenario 3 - (a) Station #13 (top left); (b) Station #15 (top right); (c) Station #24 (bottom left); (d) Combined (bottom right)

The effects of vehicle acceleration levels and percentage heavy vehicles are further illustrated in Figure 19. The figures illustrate the variation in the average flow rate at each location as a function of the average roadway density. It should be noted that the use of an average lane measure is required in order to normalize the data across the three stations given that the roadway section at Station 13 is three lanes wide while it is only two lanes wide at the other two stations. In Figure 19(a), the capacity drop is seen within the bottleneck at Station 13 (see Figure 18 for the locations of each station). The free-flow regime (traveling at an average speed of 62.5 km/hr) seems to end at a critical density of 25 veh/km/lane and a flow rate of 1320 veh/hr/lane (equivalent to a total flow of 3960 veh/hr). However, in the congested regime, a flow of 1120 veh/hr/lane (3360 veh/hr) is reached at the same critical density, which indicates a capacity drop of approximately 15 percent. This also agrees with the results in Table 1. Additionally, this finding agrees with the work in Hall and Agyemang-Duah 1991 where they

argue that the capacity drop must be measured within the bottleneck. Just downstream of the bottleneck at Station 15, there is no congestion; however the speed, as was the case at Station 13, has been significantly reduced from the free-flow speed (108 km/hr) to around 50 km/hr, as illustrated in Figure 19(b). Furthermore, the figure also demonstrates a high level of variability in the traffic stream parameters because the vehicles are accelerating. Further downstream at Station 24 in Figure 19(c), the speed increases to approximately 100 km/hr. Furthermore, the figure also demonstrates a reduction in the level of noise in the data compared to Station 15. Figure 19(d) is a combined flow-density plot showing the conditions at all three stations. It is interesting to note that the traffic stream capacity increases from approximately 2100 veh/hr/lane to 2400 veh/hr/lane between Stations 15 and 24 as vehicles group together and reach steady-state conditions.

3.6.3 Lane Reduction Bottleneck Results

The results for lane reduction bottleneck example are summarized in Table 2. They include two baseline flow conditions instead of just one condition as was the case with the on-ramp bottleneck example. This section summarizes the results from each of the four scenarios.

In Scenario 1, there are no significant flow reductions relative to Baseline 1. This indicates that the acceleration factor and the percentage of heavy vehicles are significant parameters when modeling flow reductions due to congestion. The flow reductions relative to Baseline 2 are in the range of 9 to 10 percent. According to this baseline flow condition, flow reductions are caused by the congestion upstream even when there are no heavy vehicles present and the acceleration is modeled as 100 percent of the maximum level. However, the flow reductions are further amplified by these two factors in the remaining scenarios.

In Scenario 2, the flow reductions relative to Baseline 1 are in the range of 2 to 3 percent for a total demand up to 6000-7200 vph where the flow reductions drop to around zero. The flow reductions relative to Baseline 2, which are in the range of 6 to 17 percent, represent a substantial increase as compared to Scenario 1. This indicates that by decreasing the acceleration factor from 100 to 60 percent, flow reductions increase.

In Scenario 3, the flow reductions relative to Baseline 1 are initially insignificant. However, for demands in excess of 4400 vph, the flow reductions are in the range of 1 to 7 percent which is an increase over Scenario 2. The flow reductions relative to Baseline 2 are nearly doubled from those in Scenario 2 ranging from 17 to 25 percent as compared to the 6 to 17 percent in Scenario 2. The results demonstrate that the introduction of trucks into the traffic stream produces larger discharge flow

reductions. It should be noted that for each of the four scenarios the base flow decreases as the scenario increases. Consequently, the observed increases in discharge flow reductions are not attributed to the impact of trucks but rather are caused by the lower acceleration capabilities of these heavy vehicles.

Table 2: Flow Reductions for Lane Reduction Bottleneck

	Demand Loading Case								
	1	2	3	4	5	6	7	8	9
Scenario 1: Maximum Acceleration									
Baseline 1 Flow (vph)				4292	4292	4292	4292	4292	4292
Baseline 2 Flow (vph)				4500	4800	4800	4800	4800	4800
Avg. Queue Discharge Flow (vph))				*N/A*	4286	4304	4299	4335	4372
Flow Reduction to Baseline 1 (%)				*N/A*	0.1 %	-0.3 %	-0.2 %	-1.0 %	-1.9 %
Flow Reduction to Baseline 2 (%)				*N/A*	10.7 %	10.3 %	10.4 %	9.7 %	8.9 %
Demand (vph)				4500	4800	5400	6000	6600	7200
Scenario 2: Typical Acceleration									
Baseline 1 Flow (vph)	4015	4015	4015	4015	4015	4015	4015	4015	4015
Baseline 2 Flow (vph)	4227	4178	4460	4514	4691	4691	4691	4691	4691
Avg. Queue Discharge Flow (vph)	4015	3924	3920	3927	3881	3930	4009	3995	4023
Flow Reduction to Baseline 1 (%)	*N/A*	2.3 %	2.4 %	2.2 %	3.3 %	2.1 %	0.2 %	0.5 %	-0.2 %
Flow Reduction to Baseline 2 (%)	*N/A*	6.1 %	12.1 %	13.0 %	17.3 %	16.2 %	14.6 %	14.8 %	14.2 %
Demand (vph)	4200	4300	4400	4500	4800	5400	6000	6600	7200
Scenario 3: Typical Acceleration with 6 % Trucks									
Baseline 1 Flow (vph)	3236	3236	3236	3236	3236	3236	3236	3236	3236
Baseline 2 Flow (vph)	3977	3978	3950	4007	4007	4007	4007	4007	4007
Avg. Queue Discharge Flow (vph)	3294	3247	3170	3208	3108	3111	3025	3147	3161
Flow Reduction to Baseline 1 (%)	-1.8 %	-0.3 %	2.1 %	0.9 %	4.0 %	3.9 %	6.5 %	2.8 %	2.3 %
Flow Reduction to Baseline 2 (%)	17.8 %	19.0 %	20.9 %	20.0 %	22.4 %	22.4 %	24.5 %	21.5 %	21.1 %
Demand (vph)	4000	4200	4400	4500	4800	5400	6000	6600	7200
Scenario 4: Typical Acceleration with 15 % Trucks									
Baseline 1 Flow (vph)	2552	2552	2552	2552	2552	2552	2552	2552	2552
Baseline 2 Flow (vph)	3117	3288	3124	3240	3161	3161	3161	3161	3161
Avg. Queue Discharge Flow (vph)	2608	2614	2523	2560	2515	2512	2488	2494	2490
Flow Reduction to Baseline 1 (%)	-2.2 %	-2.4 %	1.2 %	-0.3 %	1.4 %	1.6 %	2.5 %	2.3 %	2.4 %
Flow Reduction to Baseline 2 (%)	19.5 %	19.3 %	22.1 %	21.0 %	22.4 %	22.5 %	23.2 %	23.0 %	23.1 %
Demand (vph)	3600	4000	4400	4500	4800	5400	6000	6600	7200

Lastly, in Scenario 4, the flow reductions relative to Baseline 1 are not significant until a demand of 4800 vph (the base capacity of the bottleneck), then they remain fairly constant ranging from 1 to 3 percent. The discharge flow reductions caused by a 15 percent heavy vehicle composition is less than those from Scenario 3 (6 percent heavy vehicles). These findings are similar to those observed for the on-ramp bottleneck which seems to reinforce the idea that there is a “critical heavy vehicle percentage” where flow reductions are the largest. These results are consistent with findings of earlier studies [50, 51].

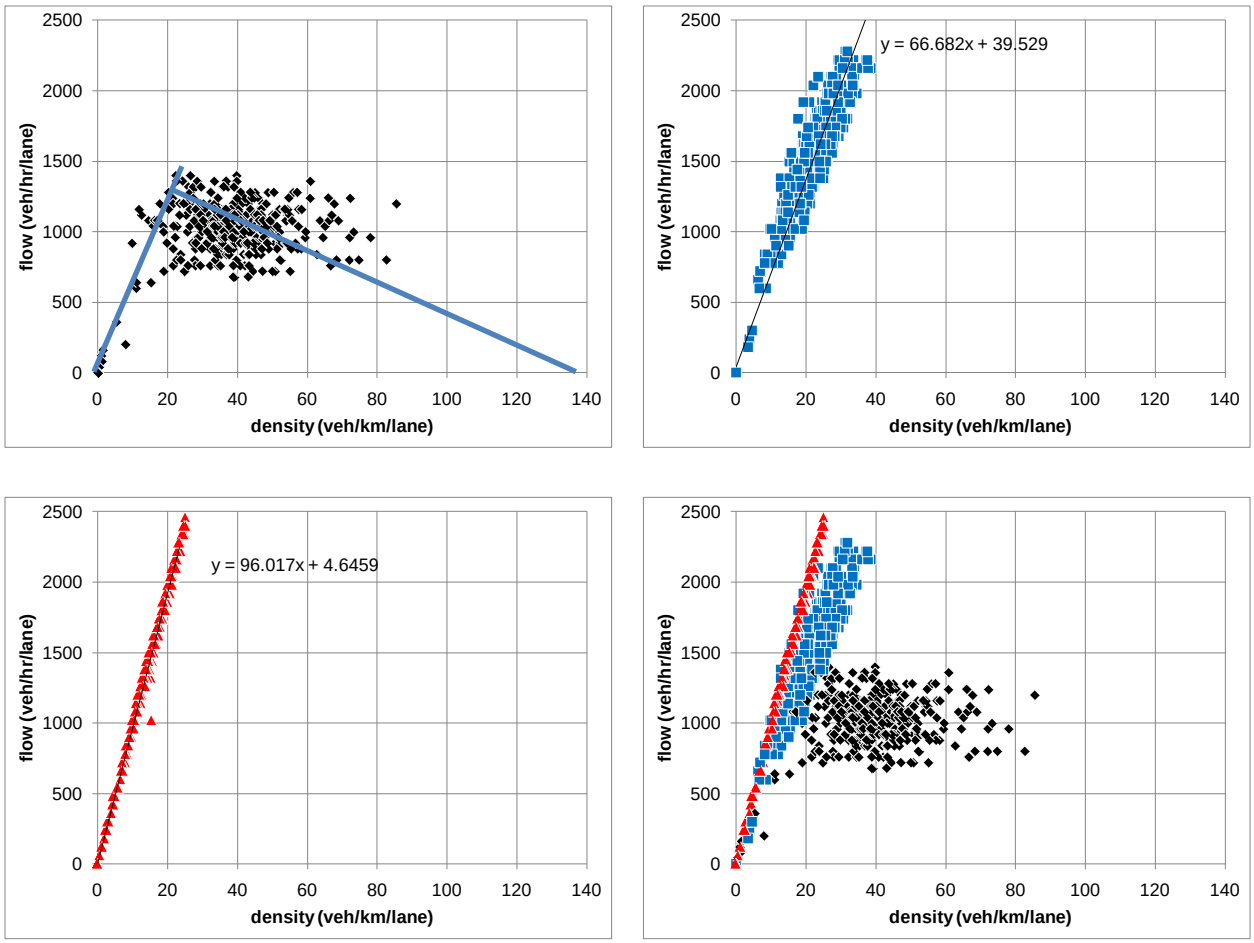


Figure 20: Flow-Density Plots for the Lane Reduction Bottleneck for Scenario 3 - (a) Station #8 (top left); (b) Station #11 (top right); (c) Station #20 (bottom left); (d) combined (bottom right)

The effects of vehicle acceleration levels and the percentage heavy vehicles are further displayed in Figure 20. In Figure 20(a), at Station 8, (see Figure 18 for the locations of each station), the free-flow regime seems to end at a critical density of 25 veh/km/lane and a flow of 1400 veh/hr/lane. In the congested regime, a flow of 1320 veh/hr/lane is reached at the same critical density, which indicates a capacity drop of approximately 6 percent. This also agrees with the results in Table 2. Just downstream of the bottleneck at Station 11 in Figure 20(b), there is no congestion however the speed has been significantly reduced from the free-flow speed (108 km/hr) to approximately 66 km/hr. Further downstream at Station 20 in Figure 19(c), the speed has increased to approximately 100 km/hr. Figure 20(d) is a combined flow-density plot showing the conditions at all three stations.

3.7 CONCLUSIONS

The central conclusion of this paper is that discharge flow reductions can be reproduced through the accurate modeling of driver acceleration behavior without the need to introduce a discontinuity in the steady-state fundamental diagram. The introduction of heavy vehicles produces further reductions in discharge flow rates. These reductions stabilize as the proportion of trucks increases. The results also demonstrate that the INTEGRATION software produces discharge flow reductions that are in the same level of magnitude as observed empirically. Furthermore, these reductions are lower for lane reduction bottlenecks in comparison to on-ramp bottlenecks. These results are consistent with empirical observations.

During the conduct of this research, further research would prove worthwhile in the following areas:

- Statistical test of significance of flow reductions from simulation runs
- Sensitivity analysis of the effect of percentage of heavy vehicles on flow reductions
- Sensitivity analysis of the effect of loop detector location both upstream and downstream of bottlenecks on flow reduction measurements

CHAPTER 4 : COMPARISON OF QUEUE DISCHARGE RATES FROM TIME-DEPENDENT AND TIME-INDEPENDENT BOTTLENECKS

Maha El-Metwally[1], Hesham Rakha[2], and Edward Chamberlayne[3]
(to be submitted for publication)

4.1 ABSTRACT

The paper introduces the concept of "capacity drop" at time-independent (e.g. freeway bottlenecks) as well as time-dependent (e.g. signalized intersection bottlenecks) bottlenecks. Using the INTEGRATION microscopic traffic simulation software the study simulates single-lane bottlenecks in order to isolate the impact of car-following behavior on the concept of capacity drop. The study demonstrates that the discharge flow rate is reduced at stationary time-independent bottlenecks after the onset of congestion. These reductions, however, are caused by a reduction in the traffic stream flow rate by moving along the steady-state fundamental diagram as opposed to a break in the fundamental diagram. Furthermore, because vehicles discharge from the same spatial location these reductions are not reduced or recovered as the traffic stream propagates downstream. Additionally, these reductions are not impacted by the level of vehicle acceleration given that they reflect the traffic stream behavior at a different location along the fundamental diagram. Alternatively, the drop in flow discharge rate caused by a time-dependent bottleneck is recoverable because vehicles discharge from a backward moving wave. These losses are demonstrated to be highly dependent on the level of acceleration that drivers are willing to exert (larger losses for less aggressive acceleration behavior). Furthermore, these losses extend over a longer distance downstream as drivers accelerate less aggressively.

4.2 INTRODUCTION

The concept of capacity is one of the most debated issues in the field of traffic flow theory. The debates about this issue not only are related to the numerical value of the capacity of different transportation facilities, but also extend to the notion of the "capacity drop" phenomenon. The concept of capacity is presented through the speed-flow relationship of the fundamental diagram, which has a parabolic shape and consists of two branches. The upper branch represents the uncongested flow regime that begins with free flow speed at low flow rates. Subsequently, as the flow rate increases, the

[1] M.S. Student, Charles E. Via, Jr. Department of Civil and Environmental Engineering, Virginia Tech
[2] Professor, Charles E. Via, Jr. Department of Civil and Environmental Engineering, Virginia Tech
[3] Ph.D. Candidate, Grado Department of Industrial & Systems Engineering, Virginia Tech

speed decreases until capacity is reached at the vertex of the parabola. On the other hand, the lower branch represents the congested flow. This branch begins when the vehicles start flowing from a zero speed and accelerating until reaching the capacity. If the speed-flow relationship is a discontinuous relationship, where capacity is higher when approached from stable flow than when approached from unstable flow, the capacity drop is presented. However, capacity drop would occur in the case of a continuous speed-flow relationship, when the flow recovering from congestion is at a lower point on the curve. In general, the capacity drop is typically occasioned by the onset of congestion.

In general, the capacity drop can be observed along uninterrupted flow facilities (e.g. freeways) as well as at interrupted flow facilities (e.g. signalized intersections). The main difference between both cases is that on freeways, congestion occurs at a stationary bottleneck which is a time-independent bottleneck. While at signalized intersections, the bottleneck is time-dependent where congestion occurs in every cycle at the onset of red. Additionally, the queue discharge point at a freeway bottleneck is fixed, as each vehicle starts accelerating from the same location within the bottleneck. On the other hand, the queue discharge point at signalized intersections is a backward moving point where each vehicle starts accelerating from zero speed from its position within the queue. Hence, further upstream vehicles reach the stop line with higher speed than the vehicles closer to the stop line.

Accordingly, the objective of this paper is to investigate the issue of capacity drop at both facility types; freeway bottlenecks and signalized intersections. This will be achieved by conducting simulation runs using the INTEGRATION microscopic traffic simulation software [28, 29] in order to evaluate and compare the flow reductions at both bottlenecks. In addition, the concept of temporary versus permanent losses will be investigated.

The remaining sections of this paper will present a background concerning the capacity drop phenomenon, an analysis of the simulation run results performed at both a freeway bottleneck and a signalized intersection, and will end with the conclusion generated by this research.

4.3 BACKGROUND

Several empirical studies have investigated the traffic conditions and causes of drops or losses in capacity along transportation networks. In one study concerning capacity analysis along freeways, Chung et al inferred that the increase in the freeway density is always followed by a drop in capacity [1]. In this study, three different types of bottlenecks on three different freeway segments were investigated to examine the relationship between vehicular density and the drop in capacity. Drops in capacity were observed at an on-ramp merge, at a freeway segment with a reduction in the number of travel lanes,

and at a horizontal curve along a freeway. The analysis of the on-ramp merge bottleneck exists on a freeway stretch of the north-bound Interstate 805 in San Diego, CA, where the bottleneck occurs due to the merge of the on-ramp at 47th St/Palm Ave with the freeway. It was concluded that a reduction of 10% in flow resulting from the formation of a queue in the freeway shoulder lane. A critical density of about 208 vehicles/km was assumed as a threshold that corresponds to the capacity drop [1]. Furthermore, the analysis of the lane reduction bottleneck took place on a freeway stretch of west-bound State Route 24 in California's San Francisco Bay Area. A capacity drop of 5% was reported, which also occurred when the density reached a critical density around 90 vehicles/km [1]. Nevertheless, what was concluded in this study [1] that specific value for critical density is causing the capacity drop is somewhat questionable. Nevertheless, generalizing a fixed critical density value to be applied at any freeway could not be accepted unconditionally, as the capacity drop can result from factors other than increased density. Contrarily, increased traffic density could occur without corresponding drop in capacity. For instance, a sudden lane changing decision from one driver can block a lane and increase its density, while the traffic volume is not high or near capacity. This discussion implies that not only high traffic density can trigger capacity drop, but it should be accompanied by high flow rates. In spite of that, no deterministic value for density of flow can accompany the reduction in capacity. In more than one study, e.g. [13, 14], Kerner investigated the flow breakdown phenomenon and agreed that the breakdown is a probabilistic phenomenon, and not deterministic as concluded by Chung et al [1].

Another study that conforms to Kerner's hypothesis was conducted by Persuad et al [3], where it was stated that the occurrence of breakdown has a probabilistic nature. To explore the probability of breakdown traffic in this study, data was collected at three different sites in Toronto, Canada using surveillance cameras and automatic collection of speed and flow data from loop detectors. It was found a capacity drop ranging from 11% to 17% in two of the sites, while in the third site the capacity drop was 26% [3]. However, this increased value in the third site was included in the analysis, without a definite explanation for it, but further work was needed [3]. Furthermore, the probability of breakdown at various traffic flow levels was investigated in the same study [3]. It was concluded that in general for the three sites, the probability of breakdown rises by increasing the flow rates. It was found that by maintaining the pre-queue flows at the same level as those that occur after a queue forms, the probability of breakdown is almost negligible. By increasing these flows 20% above the queue discharge flow, the probability of breakdown is only 10%. For more increased flows, the probability of breakdown rises dramatically. Hence, it was suggested not to operate at pre-queue exceeding the queue discharge more than 20% [3].

Furthermore, in 1990 Hall and Hall [16] investigated the effects of the formation of an upstream queue on speed and flow. Data was collected from the Queen Elizabeth Way in Ontario, Canada. It was claimed; by the authors that there is no reduction in capacity at bottlenecks downstream of a queue and that upstream queue formation had no effect on flow rates. However, this queue formation caused a reduction in observed speed [16]. Nevertheless, in 1991, Hall and Agyemang-Duah [7] came back and denied this claim, saying that Hall and Hall [16] did not have as good information with which to ascertain the time of the queue. In addition, the flows in the data they used were not heavy enough. However, in the 1991 study [7], they concluded that once queue forms upstream the bottleneck a capacity drop appears as a consequence of the way drivers accelerate away from the queue. Their conclusion was based on data collected from the same place like the previous study [16] and the results show a capacity drop of about 5 to 6% [7]. In addition to the idea of the capacity drop issue within the bottleneck, the paper argued that it should be considered first where the place to investigate the possibility of a capacity drop is [7]. The authors stated that other studies measure the capacity before and after the congestion at the location of the bottleneck, while measuring the flows at this location yields a two-branched occupancy-flow relationship [7]. The left branch represents the stable (uncongested) flow, which is higher than right branch, which represents the unstable (congested) flow. The authors argued that this station is not operating at capacity during unstable flow. Consequently, the discontinuity of the curve does not describe capacity drop. However, they concluded that the right place to measure capacity drop is the bottleneck itself [7]. Measuring at this location yields a one branch occupancy-flow relationship. This branch consists of pre-queue flows and queue discharge flows. The vertical ordinate difference between these two curves presents the capacity drop [7].

One more argument about the issue of capacity drop was raised by Banks in 1991 [18]. Even though the study concluded that bottleneck capacities decrease when flow breaks down, it stated that this applies to the capacities of individual lanes [18]. Four different bottlenecks in four sites in San Diego, CA were investigated and it was found that capacity decreased after breakdown about 10% in the left lane for Site 1. However, the capacity, when averaged across all lanes, decreases only 3%. The left lane in the other three sites had a capacity drop about 4.5% in two sites and 0.6% in the third. One the other hand, flow averaged across all lanes does not show any decrease in the three sites [18].

Saturation flow rate is defined in the 2000 HCM [4] as “The equivalent hourly rate at which previously queued vehicles can traverse an intersection approach under prevailing conditions, assuming that the green signal is available at all times and no lost times are experienced, in vehicles per hour or vehicles per hour per lane.” Saturation flow rate at signalized intersections is analogous to the concept

of capacity along freeways, as both consider 100% green time at all times. The saturation flow rate can be obtained by using its relationship with the saturation headway. The headway, in general, is the time in seconds between two successive vehicles passing the stop line, measured from the same reference point. While, the saturation headway is the stable headway between vehicles occurring after the impact of the start-up lost time vanishes. Empirical studies have shown that the headway for the first few vehicles is longer because of the start-up reaction time and the acceleration time. This headway is greatest for the first vehicle in the queue and diminishes for the following vehicles successively. The HCM assumes that the start-up lost time diminishes after the fourth vehicle and stable headways follow beginning with the fifth vehicle [4]. However, in a study by Lin and Thomas [25], they state that the queue discharge headways keep getting smaller even after the 15th stopped vehicle in the queue. As a result, the queue discharge rates continue to rise with the progress of the queue discharge. In another study by Cohen [27], queue discharges at signalized intersections using simulation analysis with application of the Pitt car-following model were analyzed. It was concluded that free-flow speed, car following parameters, vehicle length, traffic stream composition, and lane changing behavior have significant effects on queue discharge headways. In general, the study concluded that the discharge headway distribution is almost flat beyond the fifth vehicle in the queue, which is consistent to the HCM assumption.

4.4 Simulation of Stationary Time-Independent Bottlenecks

Simulation runs were made at a freeway bottleneck to investigate the capacity drop when forcing the vehicles to reduce their speed. The layout of the network used for the simulation runs is shown in Figure 21. It consists of an origin zone, two intermediate nodes, and a destination zone; the roadway segments are represented by three links: two one kilometer long one-lane links with a short fifty meter long one-lane link, in the center, that has a reduced speed limit. The vehicles travel from zone 1 to zone 2 as indicated by the arrows in the figure. The intermediate link forces the vehicles to reduce their speeds, which results in a queue forming upstream of the last link (link 3). This network configuration allows for the investigation of capacity drop without the effects of lane changing behavior. In addition, the simulation runs were made based on the network parameters shown in Table 3.

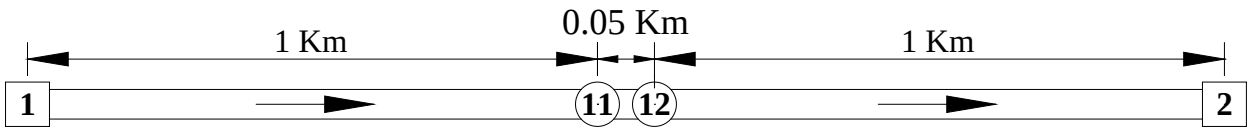


Figure 21: Network Layout for Freeway

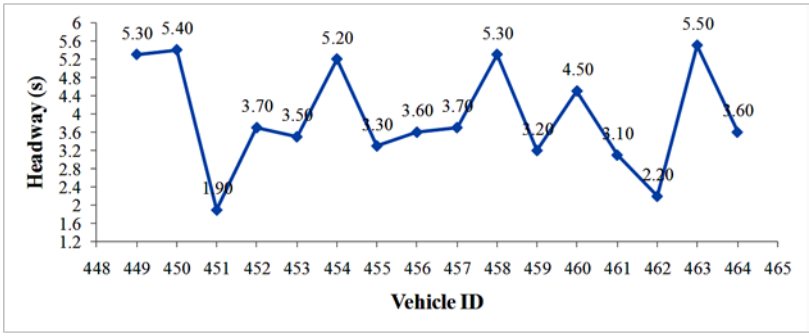
Table 3: Description of the Network and Parameters for Freeway

	Link 1	Link 2	Link 3
Length (Km)	1	0.05	1
Number of lanes	1	1	1
q_c (vphpl)	1800	1800	1800
k_j (vpkpl)	185	185	185
$u_f = u_c$ (km/h)	96	10	96

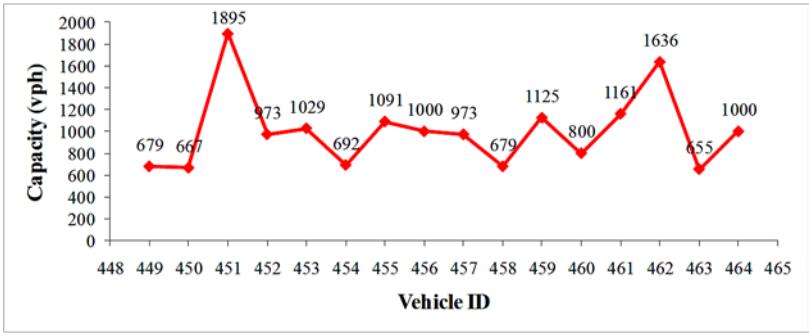
The simulation runs were made using a speed coefficient of variation of 0.05, in order to reflect differences in driver car-following behavior. Additionally, the impact of the percentage of the maximum acceleration rate was investigated by using both 100% (Scenario 1) and 60% (Scenario 2) in different simulation runs. A study by Snare 2002, found that drivers normally accelerate at 60 percent of the maximum acceleration level of their vehicle; *"Driver factors were observed that fit a normal distribution with a mean of 0.6 and a standard deviation of 0.08"* [20]. This confirmed a previous guideline published in 1954 by AASHO [21]. The simulation runs used a light duty composite vehicle with a mass of 1326 kg, a length of 4.8m, with 109kW of power.

4.4.1 Simulation Results for Freeway Scenario 1

This simulation was made with vehicles using 100% of the maximum acceleration. Fifteen vehicle IDs were tracked while they discharged from the bottleneck queue, as they were forced to reduce their speeds from free-flow speed to 10 km/h in the intermediate link. The headways for each vehicle is plotted against the vehicle ID as shown in Figure 22a. As can be seen from the figure, there does not appear to be a trend in the distribution of headways; some vehicle IDs have high headways around 5 seconds, where others have lower headways of around 2 seconds. The saturation flow rates were calculated (3600/headway) and plotted in Figure 22b.



(a) Headway

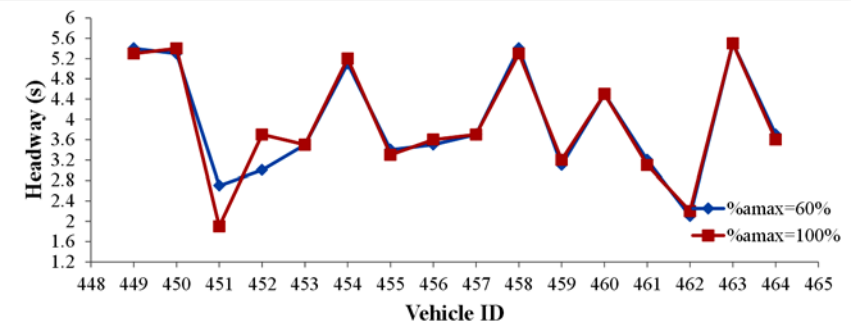


(b) Capacity

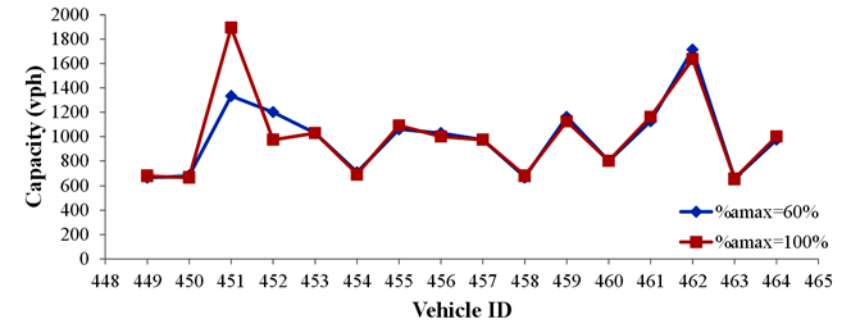
Figure 22: Distribution of Headway and Capacity for Scenario 1 for Freeway

4.4.2 Simulation Results for Freeway Scenario 2

A typical acceleration level of 60% the maximum acceleration level was simulated in this scenario. The results for the headway and the capacity for both maximum acceleration percentages, 100% and 60%, are presented in Figure 23a and Figure 23b, respectively. Visually, it can be concluded from the figures that changing the percentage of maximum acceleration has a minimum effect on the headway and flow rate results for a stationary time independent bottleneck. In addition, a statistical t-test was done to test if the results of the headways for scenario 1 and scenario 2 are significantly different. The statistical test confirmed the visual conclusion that scenario 1 and scenario 2 are not significantly different.



(a) Headway



(b) Capacity

Figure 23: Distribution of Headway and Capacity for Scenario 1 and Scenario 2 for Freeway

Furthermore, in order to better investigate if the losses in flow are recoverable or not, the time-space diagram for fifteen vehicles is plotted in Figure 24. As mentioned earlier, the free-flow speed for the first link and the third link is 96 km/h; the intermediate link forces the vehicles to reduce their speed to 10 km/h. It can be seen from the figure that at a distance of 1050 m the vehicles accelerate to increase their speed after departing the intermediate link. It can be concluded from the figure that after the vehicles depart the bottleneck and start to accelerate, the headways between the vehicles become larger, which causes a reduction in flow. This increase in the headways between the vehicles continues until the vehicles reach their destination. Hence, it can be concluded that the losses in flow resulting from a stationary bottleneck appear to be non-recoverable.

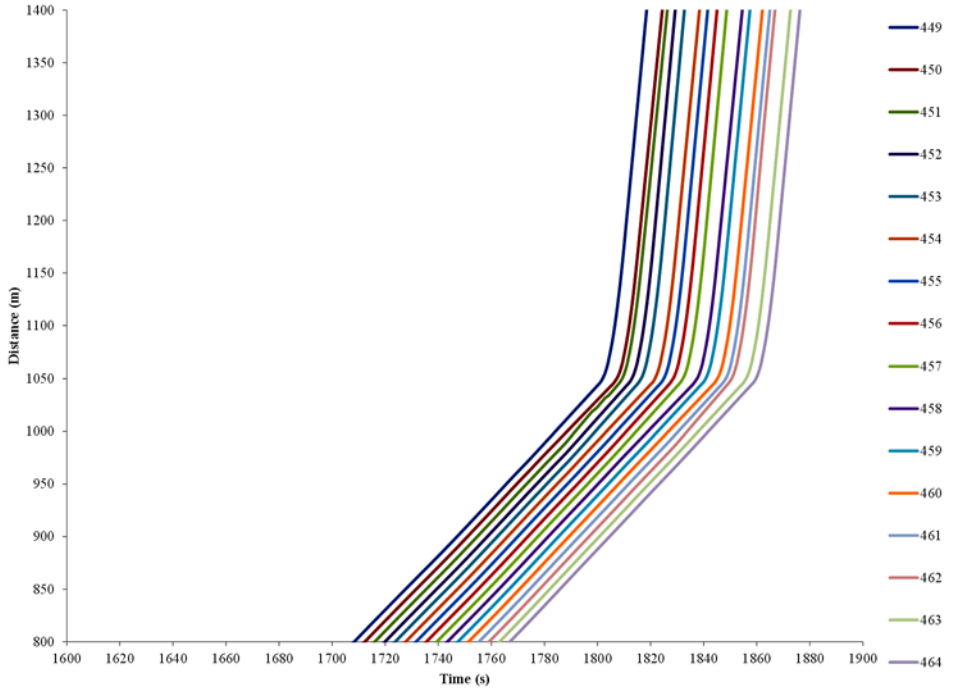
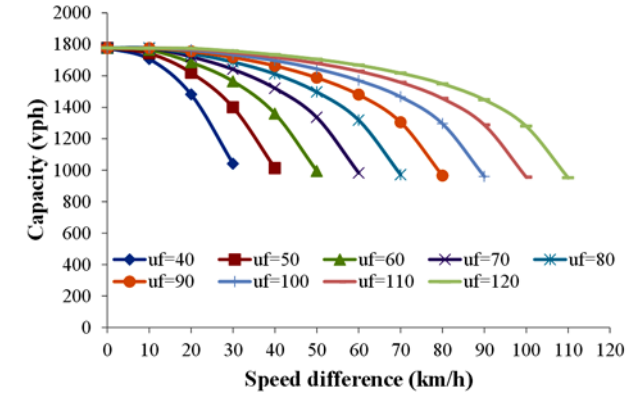


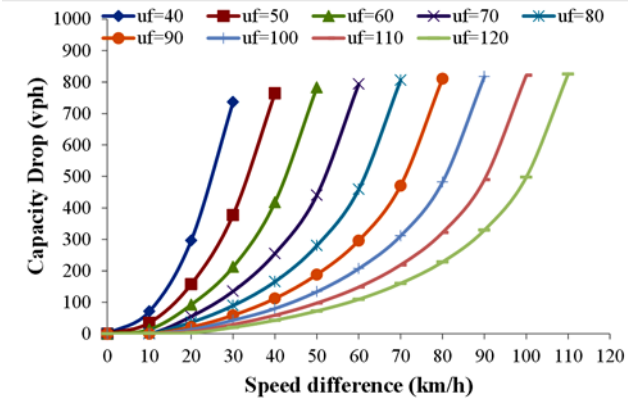
Figure 24: Time-Space Diagram for Scenario 1 for Freeway

4.4.3 Impact of Changing Speed Reduction on Capacity Drop

In both Scenario 1 and Scenario 2, the bottleneck was created by forcing the vehicles to reduce their speeds in the intermediate link only. Further simulation runs were performed in order to investigate the impact of changing the speed difference between the intermediate link and link 1 and 3 on the creation of the bottleneck. Different free-flow speeds on link 1 and 3 ranging from 40 km/h to 120 km/h at an increment of 10 km/h were investigated. For each free-flow speed scenario, a speed difference ranging from zero to the free-flow speed minus 10 km/h was investigated, again at 10 km/h intervals. The results of the capacity and capacity drop for all these speed combinations (72 combinations) were summarized in Figure 25a and Figure 25b, respectively. It should be noted that these losses mimic the reduction in flow rates along the congested regime of the steady-state fundamental diagram and are not a reflection of a break in the fundamental diagram.



(a) Capacity



(b) Capacity Drop

Figure 25: Distribution of Capacity and Capacity Drop versus Speed Difference for Different Speeds

From Figure 25a it can be concluded that by increasing the speed difference between the intermediate link and links 1 and 3, the capacity downstream the bottleneck decreases. Consequently, the capacity drop is obtained and plotted in Figure 25b. Hence, the losses in capacity increase when the speed reduction increases. In addition, the resulting losses in capacity are affected not only by the speed difference, but also by the ratio of speed difference to the free-flow speed. For the same speed difference, the higher the free-flow speed, the lower the corresponding capacity drop is.

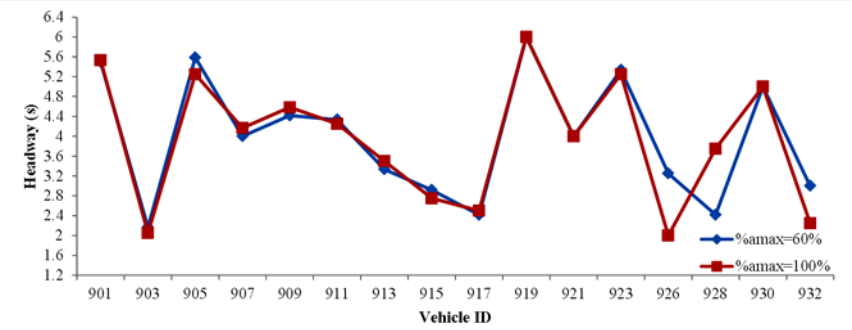
4.4.4 Simulation Results for Two-Lane Freeway for Scenario 1 and 2

After characterizing the capacity at stationary one-lane time-independent bottlenecks, similar simulation runs are made considering a two-lane time-independent bottleneck. The layout of the network and the network parameters used for the simulation are the same as was used in the case of single-lane bottleneck.

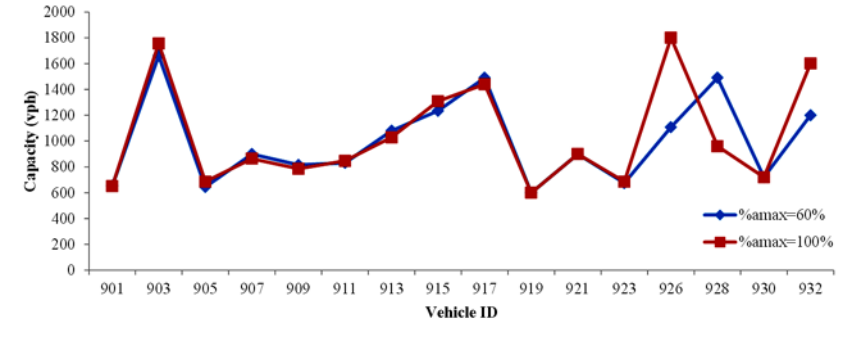
Additionally, scenario 1 (100%) and scenario 2 (60%) were used to investigate the impact of the percentage of the maximum acceleration rate on the reduction in flow discharge rate.

Fifteen vehicles were tracked while they discharged from the bottleneck queue, as they were forced to reduce their speed from the free-flow speed to 10 km/h in the intermediate link. The headways for each vehicle is plotted against the vehicle ID for both scenarios as shown in Figure 26a. As can be seen from the figure, there does not appear to be a trend in the distribution of headways; some vehicle IDs have high headways around 6 seconds, where others have lower headways of around 2 seconds. The saturation flow rates were calculated (3600/headway) and plotted in Figure 26b.

Visually, it can be concluded from the figures that changing the percentage of maximum acceleration has a minimum effect on the headway and flow rate results for a two-lane time independent bottleneck. In addition, a statistical t-test was done to test if the results of the headways for scenario 1 and scenario 2 are significantly different or not. The statistical test confirmed the visual conclusion that scenario 1 and scenario 2 are not significantly different.



(a) Headway



(b) Capacity

Figure 26: Distribution of Headway and Capacity for Scenario 1 and Scenario 2 for two-lane Freeway

Furthermore, to repeat what was done in the case of the single-lane time-independent bottleneck, the time-space diagram for fifteen vehicles is plotted in Figure 27. As illustrated in the figure, after the vehicles depart from the bottleneck, the headways between the vehicles become larger, which causes a reduction in discharge flow rate. This increase in the inter vehicle headways is caused by the fact that vehicles can only accelerate from the same spatial location. Hence, the losses in flow resulting from this bottleneck appear to be non-recoverable. Therefore, the results are consistent with the two-lane bottleneck.

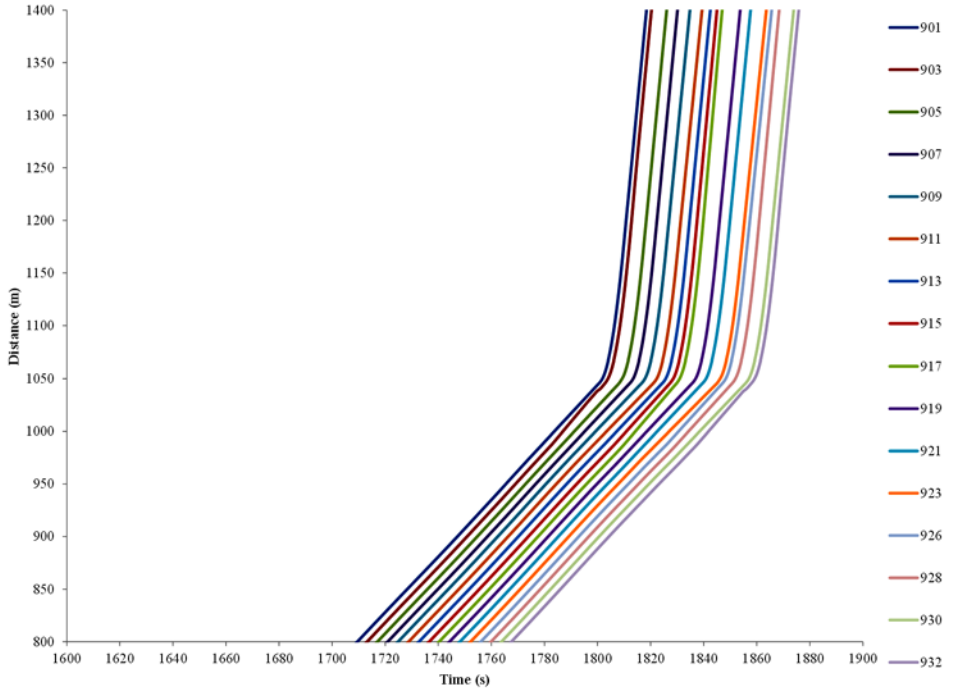


Figure 27: Time-Space Diagram for Scenario 1 for two-lane Freeway

4.5 Simulation of Time-Dependent Bottlenecks

Simulations were made also at signalized intersections in order to compare the drops in capacity to those from the freeway stationary bottleneck. Figure 28 presents the layout for the network used in these simulation runs. As shown in the figure, the network consists of an origin zone, an intermediate node, and a destination zone; the intermediate node is used to introduce a traffic signal. Accordingly, this network consists of two one-lane one kilometer long links. The reason for choosing this simple network is to focus on the start loss for the stopped vehicles in the queue without considering any other factors that could affect the results such as lane changing behavior and vehicle overtaking (passing). This is essential in order to ensure that the order of the number of vehicles in the queue is the same when

departing the stop line as the position of the vehicles in the queue is important to probe how the start loss changes from one vehicle to the next in the queue. The simulation runs were made based on the parameters in Table 4.

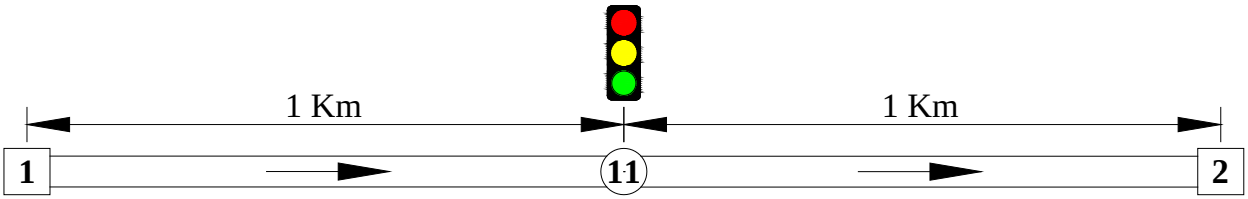


Figure 28: Network Layout for Signalized Intersection

Table 4: Description of the Network and Parameters for Signalized Intersection

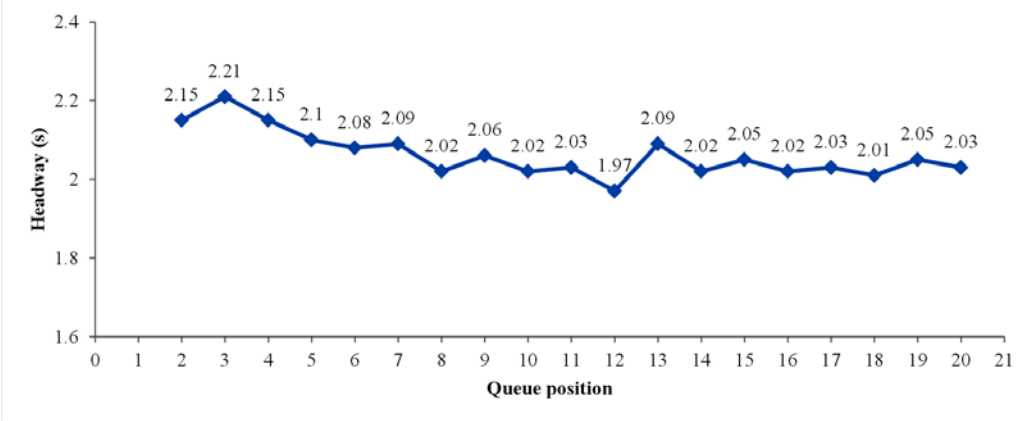
	Link 1	Link 2
Length (Km)	1	1
Number of lanes	1	1
q_c (vphpl)	1800	1800
k_j (vpkpl)	185	185
$u_f = u_c$ (km/h)	96	96

Similar to the freeway network, the simulation runs were made using a speed coefficient of variation of 0.05, light duty composite vehicles, and using a percentage of maximum acceleration of 100% (Scenario 1) and 60% (Scenario 2).

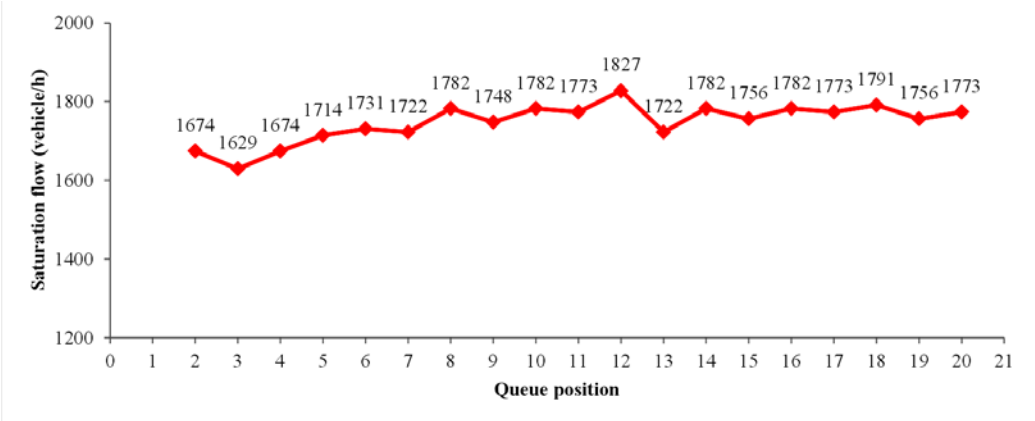
4.5.1 Simulation Results for Signalized Intersection Scenario 1

This simulation run used 100% of the maximum acceleration rate. Ten different queues were tracked while they discharged from the approach stop line at the onset of green indication. The number of queued vehicles in these queues ranged between 20 and 21 vehicles. Nevertheless, the 21st vehicle was excluded in order to ensure consistency in the number of observations for each queue position. Hence, the headway for each queued vehicle was calculated by subtracting the departure times of the leading vehicle and the following vehicle. Thereafter, the average headway for each queue position was obtained by averaging the headway for this position from the ten observed queues. Accordingly, the average headway was plotted against the queue position number, as presented in Figure 29a. It is worth mentioning that the headway was plotted starting from the second vehicle queue position as the first vehicle does not have a leading vehicle. As can be seen from the figure, the headway of the third vehicle is the largest and is 2.21 seconds and decreases almost linearly until the fifth vehicle and reaches 2.1

seconds. Starting from the fifth vehicle, the headway fluctuates around a mean of approximately 2.05 seconds. Interestingly, the headway of the second vehicle is less than the third vehicle. This could be attributed to the speed variability introduced and that the vehicle departure is random.



(a) Headway



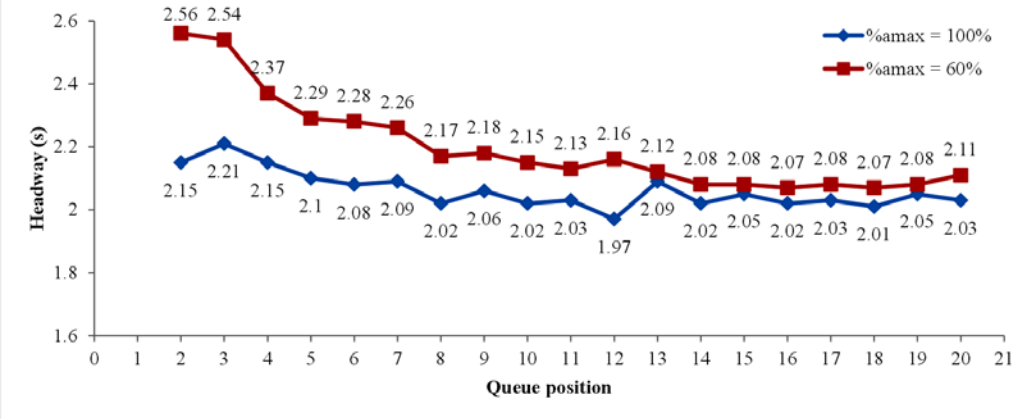
(b) Saturation Flow

Figure 29: Distribution of Headway and Saturation Flow for Scenario 1 for Signalized Intersection

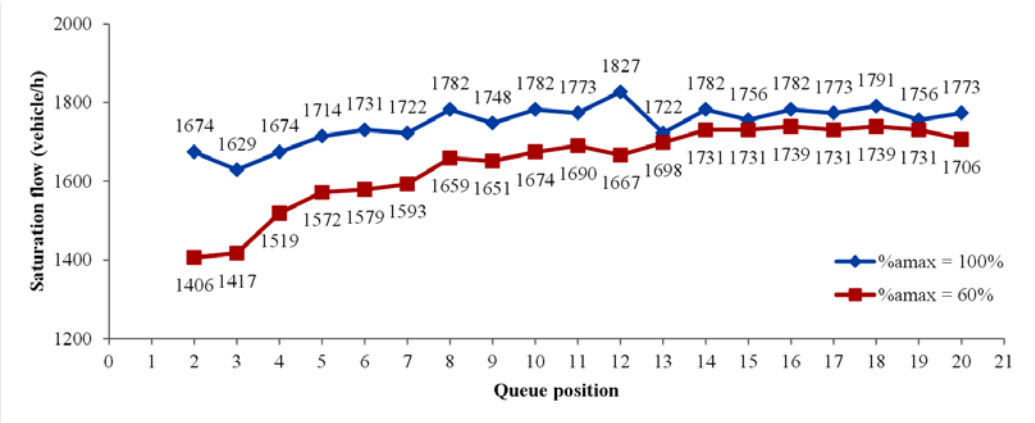
The saturation flow rate was obtained and plotted, as in Figure 29b. For the third vehicle in the queue, the saturation flow rate is smallest and is 1629 vehicles per hour. Then the saturation flow rate increases linearly until reaching the fifth vehicle, then the saturation flow rate fluctuates around 1756 vehicles per hour. These flow rates correspond to a 4% drop in saturation flow for the second vehicle, followed by 7% for the third vehicle, then the drop decreases till reaching almost a 0% drop starting from the fifth vehicle.

4.5.2 Simulation Results for Signalized Intersection Scenario 2

The impact of vehicle acceleration level on vehicle headways is examined in Scenario 2. The headways and the saturation flow rates for both maximum acceleration are presented in Figure 30a and Figure 30b, respectively. It is demonstrated from the figures that for vehicles that accelerate at typical acceleration levels incur larger headways and hence have smaller saturation flow rates when compared to vehicles that accelerate at higher acceleration levels. However, this difference in headways is highest at the beginning of the queue. In addition, the headway for the second vehicle in case of lower acceleration is the largest and not the third as in case of 100% maximum acceleration. Moreover, the headway does not stabilize even after the 20^{th} vehicle. From the results of scenario 1 and scenario 2, it can be concluded that in case of having aggressive drivers as in case of 100% acceleration level, the results are consistent with [4]. However, when having non aggressive drivers as in case of 60% maximum acceleration level, the conclusions from studies that queue discharge may not display stable saturation flow rate are also valid as in [25].



(a) Headway



(b) Saturation Flow

Figure 30: Distribution of Headway and Saturation Flow for Scenario 1 and Scenario 2 for Signalized Intersection

The time-space diagram for the twenty queued vehicles is plotted in Figure 31. From the figure it can be concluded that starting from the point that the vehicles starts to accelerate after the onset of green, the headways between the vehicles get smaller causing the saturation flow rate to return to its original value. However, these results do support the empirical findings that indicate that vehicles do not achieve steady-state conditions within four vehicles.

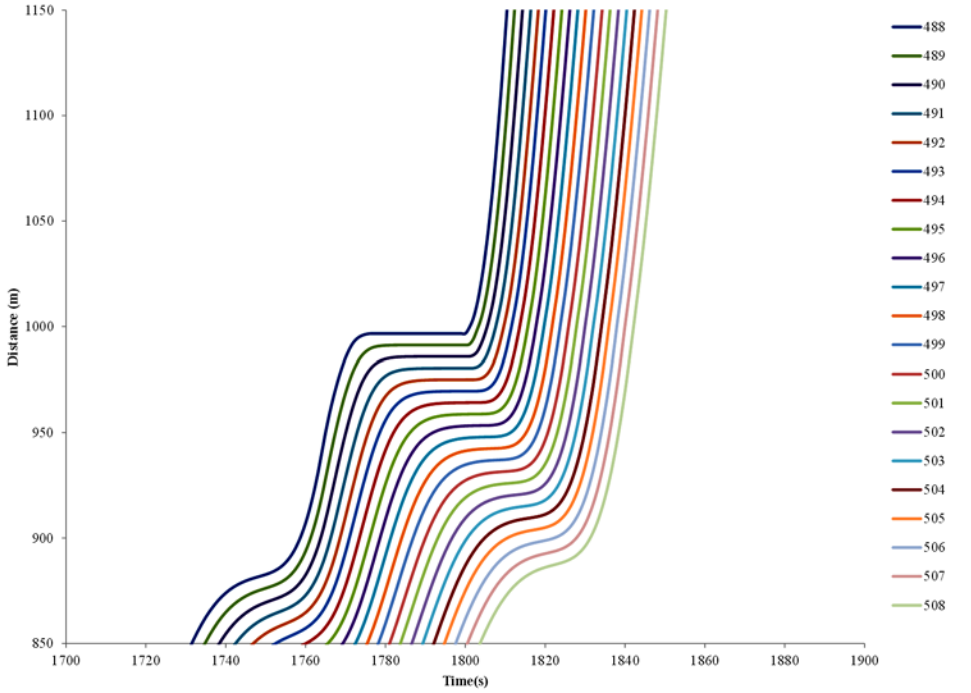


Figure 31: Time-Space Diagram for Scenario 1 for Signalized Intersection

4.6 CONCLUSIONS AND RECOMMENDATIONS

The paper introduced the concept of "capacity drop" at time-independent (e.g. freeway bottlenecks) as well as time-dependent (e.g. signalized intersection bottlenecks) bottlenecks. Using the INTEGRATION microscopic traffic simulation software the study simulated single-lane bottlenecks in order to isolate the impact of car-following behavior on the concept of capacity drop. The study demonstrated that the discharge flow rate is reduced at stationary time-independent bottlenecks after the onset of congestion. These reductions however are caused by a reduction in the traffic stream flow rate by moving along the steady-state fundamental diagram as opposed to a break in the fundamental diagram. Furthermore, because vehicles discharge from the same spatial location these reductions are not reduced or recovered as the traffic stream propagates downstream. Furthermore, these losses are not impacted by the level of vehicle acceleration given that they are a reflection of traffic stream behavior at a different location along the fundamental diagram. In addition, a two-lane stationary time-independent bottleneck was simulated, producing similar results to the single-lane bottleneck. Alternatively, the drop in flow discharge rate caused by a time-dependent bottleneck is recoverable because vehicles discharge from a backward moving recovery wave. These losses are demonstrated to be highly dependent on the level of acceleration that drivers are willing to exert (larger losses for less

aggressive acceleration behavior). Furthermore, these losses extend over a longer distance downstream as drivers accelerate less aggressively.

The results reported in this study do not account for lane-changing behavior on the reduction in flow discharge rates. A separate publication is dedicated to studying the impact of lane-changing, acceleration, and traffic stream configuration on flow discharge rates.

CHAPTER 5 : CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

5.1 CONCLUSIONS

A capacity drop occurs at the onset of congestion at freeway active bottlenecks, as well as signalized intersections at the onset of the green interval. In this thesis, the concept of capacity drop was investigated at time-independent (e.g. freeway) as well as time-dependent (e.g. signalized intersection) bottlenecks. First, it was concluded that the INTEGRATION simulation software is a valid model in simulating the reductions in traffic flow at bottlenecks, as it produces results comparable to those found in empirical studies at both merge, lane-reduction, and traffic signal time-dependent bottlenecks. Using the INTEGRATION software, it was verified that the capacity drop can be reproduced through the accurate modeling of driver acceleration behavior without the need to introduce a discontinuity in the steady-state fundamental diagram. In addition, introducing heavy vehicles in the traffic stream increases the flow reductions in the discharge flow rates.

Using the INTEGRATION simulation software, single-lane bottlenecks were simulated, in order to compare the capacity losses of time-independent and time-dependent bottlenecks without considering the impact of lane-changing behavior. It was concluded that at time-independent stationary bottlenecks, the flow reductions are not reduced or recovered as the traffic stream propagates downstream. These reductions are caused by a reduction in the traffic stream flow rate by moving along the steady-state fundamental diagram. The reason that these flow reductions are not recovered is that the vehicles discharge from the same spatial location and thus the time lost exceeds the delay associated with vehicle acceleration constraints. Consequently, these losses appear to be not affected by the level of vehicle acceleration.

At time-dependent bottlenecks, it was found that the discharge rate losses are recoverable. This is attributed to fact that the reduction in discharge flow for vehicles discharging from a backward moving recovery wave is a result of the acceleration behavior of the vehicles. Consequently, these losses are demonstrated to be highly dependent on the level of acceleration that drivers are willing to exert. These losses extend over a longer distance downstream as drivers accelerate less aggressively. Considering the lane changing on flow discharge rates at stationary bottlenecks, as is the case of on-ramp and lane-reduction bottlenecks, the flow reductions are very sensitive to the level of acceleration that drivers are willing to exert given that the bottleneck is not necessarily stationary. The cause of this

sensitivity to the level of acceleration in these bottlenecks is that lane changing allows the vehicles to start accelerating at different locations on different lanes, which is similar to moving bottlenecks. Contrarily, in case of having single lane stationary bottlenecks, all vehicles accelerate from the same location.

5.2 RECOMMENDATIONS FOR FURTHER RESEARCH

As is the case with any research effort, there still remain questions to be addressed in the future. Consequently, further research is recommended as follows:

- Investigate and identify the reasons behind having difference in the sensitivity level among different bottlenecks corresponding to vehicle acceleration.
- Attempt to produce the probabilistic nature of breakdown using simulation and comparing this behavior to empirically observed behavior.
- Develop and examine strategies and measures to reduce the impacts of the capacity drop occurrence as a result of the breakdown phenomenon. For example, one may consider vehicle-to-vehicle communication to inform vehicles of the behavior of vehicles ahead in order to prevent the breakdown in traffic flow.

  Investigate the capacity drop across the lanes in case of having a lane reduction bottleneck, in order to examine the differences of the capacity drop on lane by lane basis.

REFERENCES

1. Chung, K., J. Rudjanakanoknad, and M. J. Cassidy, *Relation between traffic density and capacity drop at three freeway bottlenecks.* Transportation Research Part B: Methodological, 2007. Vol. 41(1) p. 82-95.

2. Agyemang-Duah, K. and F. L. Hall. *Some issues regarding the numerical value of freeway capacity*. 1991. Karlsruhe, Ger: Publ by A.A. Balkema.

3. Persaud, B., S. Yagar, and R. Brownlee, *Exploration of the breakdown phenomenon in freeway traffic.* Transportation Research Record, 1998. (1634) p. 64-69.

4. *Highway Capacity Manual - HCM 2000 (Metric Units)*. 2000: Transportation Research Board.

5. Rakha, H., *Lecture Notes of CEE5604: Traffic Characteristics and Flow*. 2009, Virginia Tech.

6. Athol, P. J. and A. G. R. Bullen, *Multiple Ramp Control for a Freeway Bottleneck.* Highway Research Record, 1973. Vol. 456 p. 5-54.

7. Hall, F. L. and K. Agyemang-Duah, *Freeway capacity drop and the definition of capacity.* Transportation research record., 1991. (1320).

8. Elefteriadou, L., R. P. Roess, and W. R. McShane, *Probabilistic Nature of Breakdown at Freeway Merge Junctions.* Transportation Research Record: Journal of the Transportation Research Board, 1995. Vol. 1484 p. 80-89.

9. Lorenz, M. and L. Elefteriadou. *A Probabilistic Approach to Defining Freeway Capacity and Breakdown*. in *4th International Symposium on Highway Capacity*. 2000. Maui, HI: Transportation Research Board.

10. Roess, R. P., E. S. Prassas, and W. R. McShane, *Traffic Engineering*. 3rd ed. 2004, Upper Saddle River, N.J.: Pearson/Prentice Hall.

11. Long, G., *Start-Up Delays of Queued Vehicles.* Transportation Research Record: Journal of the Transportation Research Board, 2005. Vol. 1934(-1) p. 125-131.

12. Edie, L. C., *Car-Following and Steady-State Theory for Noncongested Traffic.* Operations Research, 1961. Vol. 9(1) p. 66-76.

13. Kerner, B. S. and S. L. Klenov, *Probabilistic breakdown phenomenon at on-ramp bottlenecks in three-phase traffic theory.* Transportation Research Record, 2006. (1965) p. 70-78.

14. Kerner, B. S., *On-ramp metering based on three-phase traffic theory downstream off-ramp and upstream on-ramp bottlenecks.* Transportation Research Record, 2008. (2088) p. 80-89.

15. Kerner, B. S., *Three-phase traffic theory and highway capacity.* Physica A: Statistical Mechanics and its Applications, 2004. Vol. 333(1-4) p. 379-440.

16. Hall, F. L. and L. M. Hall, *Capacity and speed-flow analysis of the Queen Elizabeth Way in Ontario.* Transportation research record., 1990. (1287).

17. Cassidy, M. J. and R. L. Bertini, *Some traffic features at freeway bottlenecks.* Transportation Research Part B: Methodological, 1999. Vol. 33(1) p. 25-42.

18. Banks, J. H., *Two-capacity phenomenon at freeway bottlenecks : a basis for ramp metering?* Transportation research record., 1991. (1320) p. p. 83-90.

19. Papageorgiou, M., I. Papamichail, A. D. Spiliopoulou, and A. F. Lentzakis, *Real-time merging traffic control with applications to toll plaza and work zone management.* Transportation Research Part C: Emerging Technologies, 2008. Vol. 16(5) p. 535-553.

20. Snare, M. C., *Dynamics Model for Predicting Maximum and Typical Acceleration Rates of Passenger Vehicles*, in *Civil Engineering*. 2002, Unpublished Thesis, Virginia Polytechnic Institute and State University: Blacksburg, VA. p. 133.

21. AASHO, *A Policy on Geometric Design of Rural Highways*. 1954, Washington, D.C.: American Association of State Highway Officials (AASHO).

22. Teply, S. and A. M. Jones, *Saturation flow: do we speak the same language?* Transportation research record., 1991. (1320) p. p. 144-153.

23. Akcelik, R., *Traffic Signals: Capacity and Timing Analysis*. Research Report 123. 1981, Vermont South, Vic.: Australian Road Research Board.

24. Richardson, D., B. Stephenson, J. Schnablegger, and S. Teply, *Canadian Capacity Guide for Signalized Intersections*. 1st ed. 1984, Toronto, Ontario: Institute of Transportation Engineers, District 7, Canada.

25. Lin, F.-B. and D. R. Thomas, *Headway compression during queue discharge at signalized intersections.* Transportation Research Record, 2005. (1920) p. 81-85.

26. Li, H. and P. Prevedouros, *Detailed Observations of Saturation Headways and Start-Up Lost Times.* Transportation Research Record: Journal of the Transportation Research Board, 2002. Vol. 1802(-1) p. 44-53.

27. Cohen, S. L., *Application of car-following systems to queue discharge problem at signalized intersections.* Transportation Research Record, 2002. (1802) p. 205-213.

28. M. Van Aerde & Associates, Ltd., *INTEGRATION Release 2.30 for Windows: User's Guide - Volume I: Fundamental Model Features*. 2005.

29. M. Van Aerde & Associates, Ltd., *INTEGRATION Release 2.30 for Windows: User's Guide - Volume II: Advanced Model Features*. 2005.

30. Van Aerde, M. and S. Yagar, *Dynamic integrated freeway/traffic signal networks: A routing-based modelling approach.* Transportation Research Part A: General, 1988. Vol. 22(6) p. 445-453.

31. Van Aerde, M. and S. Yagar, *Dynamic integrated freeway/traffic signal networks: Problems and proposed solutions.* Transportation Research Part A: General, 1988. Vol. 22(6) p. 435-443.

32. Dion, F., H. Rakha, and Y.-S. Kang, *Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections.* Transportation Research Part B: Methodological, 2004. Vol. 38(2) p. 99-122.

33. Rakha, H., Y.-S. Kang, and F. Dion, *Estimating Vehicle Stops at Undersaturated and Oversaturated Fixed-Time Signalized Intersections.* Transportation Research Record: Journal of the Transportation Research Board, 2001. Vol. 1776(-1) p. 128-137.

34. Rakha, H., A. Medina, H. Sin, F. Dion, M. Van Aerde, and J. Jenq, *Traffic Signal Coordination Across Jurisdictional Boundaries: Field Evaluation of Efficiency, Energy, Environmental, and Safety Impacts.* Transportation Research Record: Journal of the Transportation Research Board, 2000. Vol. 1727(-1) p. 42-51.

35. Ahn, K., H. Rakha, and A. Trani, *Microframework for Modeling of High-Emitting Vehicles.* Transportation Research Record: Journal of the Transportation Research Board, 2004. Vol. 1880(-1) p. 39-49.

36. Ahn, K., H. Rakha, A. Trani, and M. Van Aerde, *Estimating Vehicle Fuel Consumption and Emissions based on Instantaneous Speed and Acceleration Levels.* Journal of Transportation Engineering, 2002. Vol. 128(2) p. 182-190.

37. Rakha, H., K. Ahn, and A. Trani, *Development of VT-Micro model for estimating hot stabilized light duty vehicle and truck emissions.* Transportation Research Part D: Transport and Environment, 2004. Vol. 9(1) p. 49-74.

38. Rakha, H., M. Van Aerde, K. Ahn, and A. Trani, *Requirements for Evaluating Traffic Signal Control Impacts on Energy and Emissions Based on Instantaneous Speed and Acceleration Measurements.* Transportation Research Record: Journal of the Transportation Research Board, 2000. Vol. 1738(-1) p. 56-67.

39. Avgoustis, A., H. Rakha, and M. V. Aerde. *Framework for Estimating Network-wide Safety Impacts of Intelligent Transportation Systems*, in *Intelligent Transportation Systems Safety and Security Conference*, Miami, FL, 2004.

40. Van Aerde, M. *Single Regime Speed-Flow-Density Relationship for Congested and Uncongested Highways*, in *Transportation Research Board 74th Annual Meeting*, Washington, DC, 1995. 95-0802.

41. Van Aerde, M. and H. Rakha. *Multivariate Calibration of Single Regime Speed-Flow-Density Relationships*. in *6th International Vehicle Navigation and Information Systems (VNIS): A Ride into the Future*. 1995. Piscataway NJ: IEEE.

42. Demarchi, S. H. *A new formulation for Van Aerde's speed-flow-density relationship (in Portuguese)*, in *XVI Congresso De Pequisa e Ensino em Transportes*, Natal, RN, Brazil, 2002.

43. Rakha, H. and B. Crowther, *Comparison of Greenshields, Pipes, and Van Aerde Car-Following and Traffic Stream Models.* Transportation Research Record: Journal of the Transportation Research Board, 2002. Vol. 1802(-1) p. 248-262.

44. Rakha, H. *Validation of Van Aerde's Simplified Steady-State Car-Following and Traffic Stream Model*, in *Transportation Research Board 85th Annual Meeting*, Washington, DC, 2006. 01033525.

45. Mannering, F. and W. Kilareski, *Principles of Highway Engineering and Traffic Analysis*. 2nd ed. 1998, New York: Wiley.

46. Rakha, H., I. Lucic, S. H. Demarchi, J. R. Setti, and M. Van Aerde, *Vehicle Dynamics Model for Predicting Maximum Truck Acceleration Levels.* Journal of Transportation Engineering, 2001. Vol. 127(5) p. 418-425.

47. Rakha, H. and I. Lucic, *Variable Power Vehicle Dynamics Model for Estimating Truck Accelerations.* Journal of Transportation Engineering, 2002. Vol. 128(5) p. 412-419.

48. Rakha, H., M. Snare, and F. Dion, *Vehicle Dynamics Model for Estimating Maximum Light-Duty Vehicle Acceleration Levels.* Transportation Research Record: Journal of the Transportation Research Board, 2004. Vol. 1883(-1) p. 40-49.

49. Rakha, H., A. Ingle, K. Hancock, and A. Al-Kaisy, *Estimating Truck Equivalencies for Freeway Sections.* Transportation Research Record: Journal of the Transportation Research Board, 2007. Vol. 2027(-1) p. 73-84.

50. Al-Kaisy, A., J. Bhatt, and H. Rakha, *Modeling the Effect of Heavy Vehicles on Sign Occlusion at Multilane Highways.* Journal of Transportation Engineering, 2005. Vol. 131(3) p. 219-228.

51. Al-Kaisy, A., Y. Jung, and H. Rakha, *Developing Passenger Car Equivalency Factors for Heavy Vehicles during Congestion.* Journal of Transportation Engineering, 2005. Vol. 131(7) p. 514-523.