CAPACITY AND CONGESTION DUE TO INCIDENTS AT FREEWAY FACILITIES

By

CUIE LU

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2011
To my parents and my husband
ACKNOWLEDGMENTS

I am so happy that I have finished this dissertation, and am glad to thank a lot of people who have made this dissertation possible.

I am heartily thankful to my advisor, Dr. Lily Elefteriadou, for her patience, guidance, and support through these four years, for her confidence in me in conducting research. Her rich knowledge and numerous great ideas have helped me overcome many difficulties in the dissertation. I am also thankful to her for carefully reading and commenting on all my writings.

I’d like to thank my committee members Dr. Panos Pardalos, Dr. Scott Washburn, Dr. Sivaramakrishnan Srinivasan, and Dr. Yafeng Yin, for their very helpful insights and valuable suggestions to my dissertation. Thanks go out to Mr. Sampson. I have learned a lot in their classes.

This dissertation was mainly based on data collected for the National Cooperative Highway Research Program (NCHRP) 3-87 project. I would like to acknowledge Dr. Alexandra Kondyli who was responsible for data collection under that project. My thank is also extended to David Tsui of the Ontario Ministry of Transportation (MTO), Professor Baher Abdulhai of University of Toronto, Rich Bailey of Oregon Department of Transportation (ODOT), Simon Foo at McMaster University in Canada, Professor Fred L. Hall of University of Calgary, for their generous help in providing the data.

I am pleased to thank Coordinator of the Center for Multimodal Solutions for Congestion Mitigation (CMS), Ms. Ines Aviles-Spadoni, for helping me take part in various conferences and activities. She is smiling all the time and brings others pleasures. I also want to thank all my student-colleagues in transportation for their friendships and help in projects and research.
I cannot finish this dissertation without the support and encouragement of my parents and my husband. My parents’ endless love and confidence in me have given me strength and courage in pursuing the doctoral degree. Thanks go to my husband, who encourages and supports me all the time. He shares my joys and gloom in both life and study. I am so lucky that I met him at this beautiful campus when I came here.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>4</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>9</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>11</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>14</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>16</td>
</tr>
<tr>
<td>2 LITERATURE REVIEW</td>
<td>22</td>
</tr>
<tr>
<td>3 DATA COLLECTION AND METHODOLOGY</td>
<td>78</td>
</tr>
</tbody>
</table>

**ACKNOWLEDGMENTS**

**LIST OF TABLES**

**LIST OF FIGURES**

**ABSTRACT**

**CHAPTER**

1 **INTRODUCTION**

   - Background: 16
   - Objectives: 18
   - Overview of Methodology: 19
   - Organization of Dissertation: 21

2 **LITERATURE REVIEW**

   - Definition and Measurement of Freeway Capacity: 22
     - Definition of Capacity in the Highway Capacity Manual (HCM): 22
     - Definition of Capacity in Other Research: 25
     - Measurement of Capacity: 29
     - Summary: 33
   - The Process of Breakdown and Breakdown Models: 34
   - Freeway Capacity and Operation under Incident Conditions: 37
     - Impact of Incidents on Freeway Capacity: 38
     - Relationship between Incidents and Freeway Operations: 41
     - Summary: 45
   - Identification and Verification of Incidents: 46
     - Incident Detection: 46
     - Incident Prediction: 50
     - Incident Verification: 57
     - Summary: 59
   - Ramp Management Strategies Responsive to Incidents: 60
     - Ramp Closure: 61
     - Ramp Metering: 62
     - Summary: 64

3 **DATA COLLECTION AND METHODOLOGY**

   - Database Overview: 78
   - Type of Data: 78
   - Site Descriptions: 79
Methodology ........................................................................................................... 82
Data Screening ......................................................................................................... 82
Incident Verification ............................................................................................... 84
Data Analysis Procedure ....................................................................................... 86
Database Overview ............................................................................................... 87

4 QUALITATIVE ANALYSIS OF THE IMPACTS OF INCIDENTS ON OPERATIONAL CONDITIONS ................................................................. 92

Time-series Plots ..................................................................................................... 92
Density Maps ......................................................................................................... 98
Conclusions .......................................................................................................... 102

5 FREEWAY CAPACITY UNDER INCIDENT CONDITIONS .................................................. 117

Capacity under Non-incident Conditions .............................................................. 117
Capacity for Incident Conditions ........................................................................... 119
Capacity under Incident Conditions ................................................................ 119
Estimate of Capacity/Capacity Reduction under Incident Conditions ............. 123
Conclusions and Recommendations ................................................................... 125

6 AN INVESTIGATION OF THE PROBABILITY OF BREAKDOWN AND INCIDENT-INDUCED BREAKDOWN AT FREEWAYS .............................................. 133

Overview of the Product Limit Method (PLM) ....................................................... 134
Application of the Product-Limit Method for Incident Conditions ...................... 137
Probability of Demand-induced Breakdown .......................................................... 138
Description of Data and Analysis Procedure ..................................................... 139
Demand-induced Breakdown Models ................................................................. 139
Flow-based model ............................................................................................ 139
Occupancy-based model ............................................................................... 140
Speed difference-based model ..................................................................... 142
5-min standard deviation of speed-based model ......................................... 143
5-min average variation of speed-based model ............................................. 143
Probability of Incident-induced Breakdown ......................................................... 144
Process for Investigating the Probability of Incident-induced Breakdown ...... 145
Incident-induced Breakdown Models ............................................................... 145
Comparison of Demand-induced and Incident-induced Breakdown Models ...... 147
Conclusions ........................................................................................................ 148

7 FREEWAY INCIDENT DETECTION USING LIKELIHOOD OF INCIDENT FUNCTIONS ....................................................................................................... 158

Methodology ........................................................................................................ 159
Description of Data ............................................................................................ 160
Definition of Parameters and Traffic States ........................................................ 160
Calculation of the PLM ...................................................................................... 161
Results of the PLM and Incident Detection Index .......................................... 162
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Summary of the literature on capacity definition</td>
<td>66</td>
</tr>
<tr>
<td>2-2</td>
<td>Portion of freeway capacity available under incident conditions</td>
<td>67</td>
</tr>
<tr>
<td>2-3</td>
<td>Probability of lane closure due to crashes and breakdowns</td>
<td>67</td>
</tr>
<tr>
<td>2-4</td>
<td>Probability distribution of number of lanes closed</td>
<td>67</td>
</tr>
<tr>
<td>2-5</td>
<td>Capacity remaining after incidents (non-incident capacity = 1.000)</td>
<td>68</td>
</tr>
<tr>
<td>2-6</td>
<td>Percent of freeway capacity available under incident conditions</td>
<td>68</td>
</tr>
<tr>
<td>2-7</td>
<td>Performance of some common incident detection algorithms</td>
<td>69</td>
</tr>
<tr>
<td>2-8</td>
<td>Estimated binary logit model for collision incident</td>
<td>69</td>
</tr>
<tr>
<td>2-9</td>
<td>Summary of research on incident detection</td>
<td>70</td>
</tr>
<tr>
<td>2-10</td>
<td>Summary of research on incident prediction</td>
<td>71</td>
</tr>
<tr>
<td>3-1</td>
<td>Description of data available</td>
<td>88</td>
</tr>
<tr>
<td>3-2</td>
<td>Summary of data analysis</td>
<td>88</td>
</tr>
<tr>
<td>4-1</td>
<td>Relationship between incidents and beginning of congestion</td>
<td>103</td>
</tr>
<tr>
<td>4-2</td>
<td>Changes in operational conditions 5-min before/after breakdown or 5-min</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>before/after incidents</td>
<td></td>
</tr>
<tr>
<td>4-3</td>
<td>Operational conditions during different states at each data collection sites</td>
<td>104</td>
</tr>
<tr>
<td>4-4</td>
<td>Changes in density at the beginning of congestion or incidents at Toronto</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>site</td>
<td></td>
</tr>
<tr>
<td>5-1</td>
<td>Capacity estimates for non-incident conditions</td>
<td>127</td>
</tr>
<tr>
<td>5-2</td>
<td>Incident capacity and number of lanes affected</td>
<td>128</td>
</tr>
<tr>
<td>5-3</td>
<td>Comparison of incident capacity at different sites</td>
<td>128</td>
</tr>
<tr>
<td>5-4</td>
<td>Comparison of percent of freeway capacity available under incident conditions</td>
<td>129</td>
</tr>
<tr>
<td>5-5</td>
<td>Effects of various factors on parameters of incident capacity</td>
<td>130</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5-6</td>
<td>Estimates of parameters for regression to estimate the minimum 10-min flow rate</td>
<td>130</td>
</tr>
<tr>
<td>5-7</td>
<td>Estimates of parameters for regression to estimate the total capacity reduction</td>
<td>131</td>
</tr>
<tr>
<td>6-1</td>
<td>Summary of site characteristics and data available</td>
<td>151</td>
</tr>
<tr>
<td>7-1</td>
<td>Overview of site characteristics and data</td>
<td>171</td>
</tr>
<tr>
<td>7-2</td>
<td>Relationship between likelihood of incident and each parameter for non-congested conditions</td>
<td>171</td>
</tr>
<tr>
<td>7-3</td>
<td>Relationship between likelihood of incident and each parameter for congested conditions</td>
<td>171</td>
</tr>
<tr>
<td>7-4</td>
<td>Characteristics of evaluation data and evaluation results</td>
<td>172</td>
</tr>
<tr>
<td>7-5</td>
<td>Comparison results of the proposed index to previous algorithms</td>
<td>172</td>
</tr>
<tr>
<td>A-1</td>
<td>Format of data at I-15 SB</td>
<td>184</td>
</tr>
<tr>
<td>A-2</td>
<td>Format of data at I-5 NB</td>
<td>184</td>
</tr>
<tr>
<td>A-3</td>
<td>Format of data at the Queen Elizabeth Way (QEW) site</td>
<td>184</td>
</tr>
<tr>
<td>A-4</td>
<td>Format of data at the US 217 SB.</td>
<td>185</td>
</tr>
<tr>
<td>A-5</td>
<td>Format of data at the US 494 EB.</td>
<td>185</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Flow chart of the dissertation</td>
<td>21</td>
</tr>
<tr>
<td>2-1</td>
<td>Illustration of three parameters on time-series plot of flow and speed</td>
<td>72</td>
</tr>
<tr>
<td>2-2</td>
<td>Probability of F-S transition and the corresponding capacity</td>
<td>72</td>
</tr>
<tr>
<td>2-3</td>
<td>Capacity distributions for a 3-lane freeway with and without variable speed control (13.5% average truck percentage, 5-minute interval)</td>
<td>73</td>
</tr>
<tr>
<td>2-4</td>
<td>Probability of breakdowns in 15 min by Elefteriadou et al., reproduced</td>
<td>73</td>
</tr>
<tr>
<td>2-5</td>
<td>Probability of breakdown versus observed flow rate - Site “A”</td>
<td>74</td>
</tr>
<tr>
<td>2-6</td>
<td>Relationship between incident types and number of lanes affected</td>
<td>74</td>
</tr>
<tr>
<td>2-7</td>
<td>Relationship between v/c and accident rate at basic freeway section</td>
<td>75</td>
</tr>
<tr>
<td>2-8</td>
<td>Structure of basic California incident detection algorithm</td>
<td>75</td>
</tr>
<tr>
<td>2-9</td>
<td>Predicted collision likelihood</td>
<td>76</td>
</tr>
<tr>
<td>2-10</td>
<td>Example of incident occupancy</td>
<td>76</td>
</tr>
<tr>
<td>2-11</td>
<td>Incident-caused occupancy at multiple stations</td>
<td>77</td>
</tr>
<tr>
<td>3-1</td>
<td>Location of detectors at I-15 SB</td>
<td>89</td>
</tr>
<tr>
<td>3-2</td>
<td>Location of detectors at I-5 NB</td>
<td>89</td>
</tr>
<tr>
<td>3-3</td>
<td>Schematic of detectors at the Queen Elizabeth Way (QEW) Site</td>
<td>89</td>
</tr>
<tr>
<td>3-4</td>
<td>Location of detectors at QEW</td>
<td>90</td>
</tr>
<tr>
<td>3-5</td>
<td>Schematic of the Portland site</td>
<td>90</td>
</tr>
<tr>
<td>3-6</td>
<td>Location of detectors at the Portland site</td>
<td>90</td>
</tr>
<tr>
<td>3-7</td>
<td>Location of detectors at the Minneapolis site</td>
<td>91</td>
</tr>
<tr>
<td>3-8</td>
<td>Verification of incident on Jan 28, 2005</td>
<td>91</td>
</tr>
<tr>
<td>4-1</td>
<td>Time series speed and flow plots for condition with no incident.</td>
<td>105</td>
</tr>
<tr>
<td>4-2</td>
<td>Time series plots for incidents occurring before breakdown (March 9, 2005)</td>
<td>106</td>
</tr>
</tbody>
</table>
4-3 Time series plots for incident during the congested period (Oct 6, 2005) ....... 107
4-4 Time series plots for incident occurring downstream condition (Sep 19, 2005) 108
4-5 Time series speed and flow plots for minor incidents (Nov 8, 2005) ............ 109
4-6 Distributions of the changes in operations for no incident and incident conditions. .................................................................................................................. 110
4-7 Density map of Feb 10 2005 without incident ............................................. 111
4-8 Density map of March 9 2005 with incident at 17:17-18:01 at 490DES .......... 112
4-9 Density map of May 25 2005 with incident at 9:57-10:27 at 440DES .......... 113
4-10 Density map of Oct 6 2005 with incident at 8:02-8:38 at 440DES (5-min) .... 114
4-11 Density map of Oct 6 2005 with incident at 8:02-8:38 at 440DES (1-min) .... 115
4-12 Density Map of Sep 22 2005 with Incident at 17:16-19:46 at 510DES ....... 116
5-1 Capacity parameters by number of lanes .................................................. 132
6-1 Probability of demand-induced breakdown based on flow ...................... 152
6-2 Probability of demand-induced breakdown based on occupancy .......... 152
6-3 Probability of demand-induced breakdown at the two bottleneck locations at Toronto site .................................................................................................................. 153
6-4 Probability of demand-induced breakdown based on normalized speed difference .................................................................................................................. 154
6-5 Probability of demand-induced breakdown based on 5-min std.v ............. 154
6-6 Probability of demand-induced breakdown based on 5-min cvs .............. 155
6-7 Probability of incident-induced breakdown ............................................. 156
6-8 Comparison of probabilities of demand-induced breakdown and incident-induced breakdown .......................................................... 157
7-1 Likelihood of incident for non-congested conditions ................................. 173
7-2 Likelihood of incident for congested conditions ....................................... 174
7-3 Likelihood of incident potential based on each parameter for incidents occurring before congestion .......................................................... 175
7-4  Likelihood of incident potential based on each parameter for incidents occurring during congestion.......................................................... 176

7-5  Evaluation results of incident detection ................................................................. 179
Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CAPACITY AND CONGESTION DUE TO INCIDENTS AT FREEWAY FACILITIES

By

Cuie Lu

December 2011

Chair: Lily Elefteriadou
Major: Civil Engineering

Incidents lower freeway capacity and increase delay, especially during high
demand and oversaturated flow conditions. This dissertation investigates freeway
capacity and congestion due to incidents through four aspects that are detailed below,
based on data collected from five freeway facilities in North America.

A qualitative analysis of the impact of incidents on operational conditions was first
conducted by constructing time-series plots and density maps. It was determined that
incidents which are verified to affect traffic flow are mostly collisions with durations
greater than 10 minutes. Changes in flow, speed and occupancy at the beginning of
congestion that was caused by incidents are much steeper than that at the beginning of
recurrent (i.e., demand-induced) congestion. Generally, the speed of shockwaves for
incident-induced breakdowns is larger than that for demand-induced breakdowns.

Next, maximum throughput-related values were obtained to estimate capacity
under non-incident conditions as well as under incident conditions. Under non-incident
conditions, the data indicate that three-lane freeways are the most efficient in terms of
per lane capacity. Two multiple linear regression models were developed to predicts the
capacity of a section during incident conditions and capacity reduction during incidents.
The results suggest that apart from the number of lanes affected by incidents, incident category and total number of lanes are significant in incident capacity estimation.

The Product Limit Method (PLM) is reviewed and extended to estimate the probability of breakdown caused by incidents (defined as incident-induced breakdown) based on five different traffic parameters. The results show that flow is more appropriate than the other parameters in estimating the probability of breakdown. The results are also compared to the probability of demand-induced breakdown based on the same parameters, and found that at the same flow rate, the probability of incident-induced breakdown is higher than the probability of demand-induced breakdown.

Finally, the dissertation investigates whether the incident occurrence (primarily crashes) can be detected based on likelihood of incident functions using the PLM. The analysis considers seven different traffic parameters. The likelihood functions of incident based on each parameter is first developed using the PLM for congested and non-congested conditions at each site. An incident detection index is then proposed based on the PLM results. The analysis results show that, at all the sites, four parameters are significant in incident detection: flow, occupancy, standardized speed difference, and 5-min average variation of speed. The proposed index is evaluated with detector data (both clean data and raw data), and compared to previously developed algorithms. The results show that the proposed index yields higher detection rate and lower mean detection time. The clean data generate fewer false alarms than the raw data. It is also found that most of the false alarms occur during congested conditions. The false alarm rate can be reduced by decreasing the index coefficients; however, the detection rate would also be reduced.
CHAPTER 1
INTRODUCTION

Background

Freeway incidents are nonrecurring unexpected events that might disrupt traffic flow. These events include crashes, vehicle breakdowns, etc, which lead to delays, wasted fuel and lost productivity. According to Farradyne (2000), 70% of all incidents are reported, among which 80% are disabled vehicles, 10% are crashes and 10% are other types of incidents. According to their study, among the crashes, 60% occur on the shoulder with a duration of 45 - 60 minutes and 40% block lanes with a duration of 45 - 90 minutes. Among the others, 70% occur on the shoulder with a duration of 15 - 30 minutes and 30% block lanes with a duration of 30 - 45 minutes. Incidents account for approximately 60% of all urban freeways delay in US (Federal Highway Administration, 2000).

At low and medium traffic demand conditions, incidents may lower the travel speed at the incident location and thus increase delay. The impact of incidents during high demand and oversaturated freeway flow conditions is much larger than that during low and medium traffic demand conditions due to two reasons. First, as indicated by Persaud and Azbik (1993), incidents are more likely to occur during congested operations than during uncongested operations at similar flows on the same highway. Second, traffic flow at higher density is relatively unstable and thus more easily affected by other disruptions such as incidents. Therefore, studying the impacts of incidents on freeway operation for high demand and oversaturated conditions is very important for freeway operation efficiency and safety.
Incidents may lead to capacity reduction, as they can block one or more lanes and/or the adjacent shoulder. Previous research, such as the Highway Capacity Manual (HCM, 2010) and that by Goolsby (1971), estimated the capacity during an incident based on the number of lanes closed by the incident, and reported this capacity reduction as a deterministic value. However, there is no clear relationship between freeway capacity and other incident characteristics (e.g., incident duration), or between freeway capacity and geometric characteristics. Nor is there estimated capacity during an incident based on a large dataset. Moreover, it is not clear whether the capacity value during an incident is a constant. Therefore, it is necessary to estimate capacity for incident conditions based on some other incident characteristics and geometric conditions, all of which are based on a larger dataset. This could help estimate capacity during incidents more accurately.

The breakdown of flow on freeways is usually defined as the beginning of congestion. It describes the transition from relatively free-flowing traffic, at speeds in the vicinity of the speed limit, to congestion, usually defined as stop-and-go traffic. While several breakdown models have been proposed, such as the probabilistic model by Elefteriadou et al (1995) and the model based on Markov process by Evans et al. (2001), there is limited research on the phenomenon of freeway breakdown induced by incident, which will be defined as incident-induced breakdown in this dissertation. When an incident occurs, the freeway capacity at that section may be reduced, and vehicles may accumulate upstream. The mechanism of the transition stage in which flow is changed from uncongested conditions to congested conditions when an incident occurs
is not clear. It is also unknown whether an incident-induced breakdown model could be
developed. Thus, further research is deserved in these areas.

With respect to the impacts of incidents on freeway operation, various incident
management programs have been proposed and developed. Incident detection is a
crucial step in incident management. Prompt and reliable incident detection is vital in
reducing incident congestion and secondary incidents. However, existing incident
detection algorithms do not consider many variables and have difficulty in differentiating
incidents from congestion. Previous research has shown that ramp metering has the
potential to delay or prevent breakdown (Elefteriadou et al., 2009), and also the
flexibility to be responsive during the incidents by updating the capacity value (Chang et
al., 1994, Bogenberger et al., 2001). However, no incident probability model is used in
those ramp metering algorithms.

This dissertation reports the results of the study on the impact of incidents on
freeway flow from three aspects: impact on operational conditions, impact on capacity,
and impact on congestion, and propose some recommendations such as ramp metering
in response to incidents.

**Objectives**

The objectives of this dissertation are specified below.

- Conduct a qualitative analysis of the impacts of incidents on operations. The
effects of incidents on operations will be studied by comparing flow, speed and
occupancy as well as their variances before, during and after the incidents. The
characteristics of breakdown and incident-induced breakdown will also be
compared. This would provide general information about the effects of incidents on
freeway capacity, breakdown patterns, and incident detection.

- Estimate freeway capacity under incident conditions. The relationship between
capacity and incident characteristics, as well as the variability of capacity during
different incident conditions, will be explored. The results of this research may be
incorporated into the HCM to provide a more accurate estimation of capacity for
different incident conditions. They can also be used in incident management programs and ramp metering strategies, to improve throughput and safety under incident conditions. Moreover, the results can be applied to planning models to take into consideration various incident scenarios.

- Investigate the probability of incident-induced breakdown. The potential of developing a probabilistic model of incident-induced breakdown will be investigated, based on previous research on breakdown and field data analysis of traffic flow during incidents. Such a model could be used in traffic management and operations of freeway facilities (for example in ramp metering and variable speed limit algorithms.)

- Investigate the relationship between incident probability and operational conditions. This objective is based on the observation that operational conditions are distinct under non-incident and various incident conditions. Such a relationship can be incorporated into ramp metering and other freeway management tools, to increase throughput and decrease delay on the freeway.

The flow chart of this dissertation is described in Figure 1-1.

**Overview of Methodology**

This dissertation focuses on exploring the relationship between incidents, capacity, and breakdown. To achieve these goals, previous research was reviewed, and the approach defined for this dissertation was developed considering these findings. The analysis will be based on a large dataset, which includes traffic data, incident data, weather data and geometric data that were collected from several freeway sections in North America. Based on the dataset, the following tasks were completed to achieve each of the four objectives in this dissertation:

Task 1: Literature Review. Previous research related to the definition and measurements of freeway capacity under various conditions, the process and models of flow breakdown, the impact of incidents on freeway operation and capacity, incident verification and identification, as well as ramp metering strategies responsive to incidents, are reviewed.
Task 2: Data Assembly. The data used are based on the National Cooperative Highway Research Program (NCHRP) 3-87 project (ELEFTERIADOU et al., 2009). Traffic data, weather data, incident data, and geometric data collected from five freeway sections for the project are used in this dissertation. Additional data are collected later.

Task 3: Data Analysis. The data are firstly screened by weather and incident conditions. Only data with good weather conditions are kept for further analysis. Then incidents were verified by time, location, and effect, and further separated into different groups according to the time and location of incidents as well as the time of congestion. Based on the analysis and literature review, an analysis method to achieve the objectives was developed.

Task 4: Explore the Relationship between Incidents and Operational Conditions. The relationships between incidents and beginning of congestion are investigated through time series plots and density maps. Operational conditions such as flow, speed and occupancy as well as their changes before, during and after the incidents are also obtained. This provides general information about the effects of incidents on freeway capacity, congestion, and incident prediction.

Task 5: Define and Measure Freeway Capacity for Incident Conditions. Based on the data analysis results, a definition of freeway capacity for incident conditions is proposed and a method to measure the capacity for incident conditions is specified.

Task 6: Investigate the Relationship between Incident-induced Breakdown Probability and Incidents. The potential of developing a probabilistic model of incident-induced breakdown is investigated, and the product limit method is considered and used in developing the model.
Task 7: Detect Incidents as a Function of Operational Conditions. The relationship between incident probability and operational conditions is studied through the product limit method.

**Organization of Dissertation**

The second chapter reviews previous literature and provides conclusions on the definition and measurement of capacity, the process and models of breakdown, the impact of incidents on freeway operation and capacity, incident identification and verification, and ramp management strategies responsive to incidents. The third chapter overviews the database and describes the methodology. The fourth chapter qualitatively analyzes the impacts of incidents on operational conditions. The fifth chapter estimates freeway capacity for both non-incident and incident conditions. The sixth chapter investigates the potential of developing an incident-induced breakdown probability model. The seventh chapter investigates the relationship between operational conditions and incident occurrence. The last chapter concludes the dissertation and presents recommendations.

![Flow chart of the dissertation](image)

Figure 1-1. Flow chart of the dissertation
CHAPTER 2
LITERATURE REVIEW

This chapter reviews previous research related to freeway capacity, flow breakdown and incidents. The contents include the definition and measurement of capacity, the process and models of flow breakdown, the impacts of incidents on freeway operations and capacity, incident identification and verification, as well as ramp metering strategies responsive to incidents.

Definition and Measurement of Freeway Capacity

A main domain in incident research is capacity. The term “Capacity” is used to quantify the traffic-carrying ability of transportation facilities. The capacity value is used in designing or rehabilitating highway facilities. It is also used in evaluating whether an existing facility can satisfy the expected traffic demand. The definition and measurement of freeway capacity have been studied extensively, yet no agreement has been achieved. This section reviews literature related to the definition and measurement of freeway capacity for different traffic conditions.

Definition of Capacity in the Highway Capacity Manual (HCM)

Originally published in 1950, the HCM was the first and most often used document to quantify the concept of capacity for transportation facilities.

The HCM 1950 (HCM, 1950) defined three levels of roadway capacity: basic capacity, possible capacity and practical capacity. Basic capacity was “the maximum number of passenger cars that can pass a point on a lane or roadway during one hour under the most nearly ideal roadway and traffic conditions which can possibly be attained.” Possible capacity was “the maximum number of vehicles that can pass a point on a lane or roadway during one hour under prevailing roadway and traffic
conditions.” Practical capacity was a lower selected volume that is meant to avoid high traffic density and improve drivers' freedom of maneuver. Among the three definitions, the “possible capacity” was more similar to the definition used today with the exception that it didn’t consider control conditions. It had also realized the variability of capacity by considering the roadway and traffic conditions, which may vary, in the definition. The “practical capacity” was related to the level of service concept and the service volume that corresponding to a specific LOS.

The HCM 1965 (HCM, 1965) defined capacity as the maximum number of vehicles which had a reasonable expectation of passing over a given section of a lane or roadway in one direction (or in two lane direction for a two-lane or three-lane highway) during a given time period under prevailing roadway and traffic conditions, similarly to the “possible capacity” definition of the HCM 1950. This edition of the manual changed “basic capacity” to be “capacity under ideal conditions”, and replaced “practical capacity” to be “service volume under a series of conditions”.

The HCM 1985 (HCM, 1985) defined the capacity of a facility as the maximum hourly flow rate at which persons or vehicles reasonably can be expected to traverse a point or a uniform section of a lane or roadway during a given time period under prevailing roadway, traffic, and control conditions. It stressed that capacity refers to a rate of vehicular or person flow during a specified period of interest, which was most often a 15-minute period. This definition recognized the potential for substantial variations in flow during an hour, and focused analysis on intervals of maximum flow.

The HCM 2000 (HCM, 2000) defined the capacity of a facility as the maximum hourly rate at which persons or vehicles reasonably can be expected to traverse a point
or a uniform section of a lane or roadway during a given time period under prevailing roadway, traffic, and control conditions. However, it still defined capacity as a deterministic value, although it stated that capacity was not the absolute maximum flow rate observed on such a facility, as driver characteristics vary from region to region, and thus the absolute maximum flow rate can vary from day to day and from location to location.

The HCM 2010 (HCM, 2010) defines the capacity of a freeway facility as the capacity of the critical segment (breakdown first) among those segments composing the default facility. It indicates that it is important to evaluate the individual segment demands and capacity.

For a long time, traffic engineers have realized the inadequacy of the capacity definition mainly for two reasons (Elefteriadou, 2004): firstly, the expression “maximum hourly rate…that can reasonably be expected to...” was not specific enough to obtain an estimate of capacity from field data, secondly, many previous research studies had shown that the maximum throughput varies from day to day and from location to location. Additionally, the HCM 2000 stated that the “turbulence due to merging and diverging maneuvers does not affect the capacity of the roadways involved.” However, observations show that capacities at different links are different (Jia et al, 2001). Therefore, there is a desire to define capacity under different conditions. The latest HCM pointed out this issue and suggested that freeway capacity should be measured at the critical segment (bottleneck).

In summary, the definition of capacity within the HCM has evolved overtime, and the capacity value provided by it has increased over time from 2000 pc/hr/ln to 2400.
pc/hr/ln for a basic freeway segment. There has been a suggestion to consider the variability of maximum volumes in the capacity definition.

**Definition of Capacity in Other Research**

In parallel to the HCM, there have been other research efforts that study on the definition of capacity. The following reviews references related to the definition of capacity in other studies, especially capacity at freeway bottlenecks, which is defined as the freeway segment after the merge point here, that is, the end of the acceleration lane.

Agymang-Duah and Hall (1991) plotted the maximum pre-queue flows and mean queue discharge flows in 15-minute intervals and found the two distributions are similar. Based on the mean values of the two observed flows, they recommended 2300 pc/h/ln as the capacity under stable flow conditions and 2200 pc/h/ln for the post-breakdown conditions. However, the effect of site characteristics on capacity was not considered in their study. Similarly, Persaud and Hurdle (1991) suggested the mean discharge flow as the most appropriate way to define capacity. This conclusion was drawn partly due to the consistency the researchers observed in its day-to-day measurement.

Minderhoud et al. (1997) researched empirical capacity estimation methods for uninterrupted flow facilities and recommended the product limit method. According to the method, non-congested flow data were used to estimate the capacity distribution. However, this paper did not discuss the measurement of discharge flow.

Cassidy and Bertini (1999) analyzed data from two freeway bottlenecks in Toronto, Canada. They observed lower discharge flows than the flows measured prior to the queue formation, and found that the discharge flow rates remain consistent from day to day while the maximum pre-queue flows were generally unstable during short periods.
Thus, the authors suggested using the long-run queue discharge flow as the bottleneck capacity, due to its reproducibility. However, defining the reproducible discharge flow as capacity somewhat ignores the random nature of capacity.

Freeway capacity is also related to flow breakdown, which is the transition between proper operation and non-acceptable flow conditions. Lorenz and Elefteriadou (2001) observed that flow rates may remain constant or even increase after breakdown, and that the flow drop after breakdown might be contingent upon the particular flow rate at which the facility breaks down. Thus, they suggested that a probability component should be incorporated into the freeway capacity definition. The proposed freeway capacity was the rate of flow (in pcphpl for a specified time interval) corresponding to the expected probability of breakdown under prevailing conditions. However, this paper did not discuss maximum pre-breakdown flow nor did it compare the breakdown flows to maximum discharge flows for each observation day.

Elefteriadou and Lertworawanich (2003) defined and examined three flow parameters: the breakdown flow, the maximum pre-breakdown flow, and the maximum discharge flow, based on freeway traffic data at two sites over a period of several days. The breakdown flow was defined as the 5-min flow (or 15-min flow) immediately prior to the breakdown event. The maximum pre-breakdown flow was the maximum flow observed prior to the occurrence of congestion. And the maximum discharge flow was the maximum flow observed after the breakdown occurrence and prior to the recovery to non-congested conditions. Figure 2-1 illustrates these terms. The authors concluded that the three parameters vary on a relatively large range and follow the normal distribution, and that the value of breakdown flow is almost always lower than both the
maximum pre-breakdown flow and the maximum discharge flow. Moreover, they found that the maximum pre-breakdown flow tends to be higher than the maximum discharge flow at one site but the opposite occurs at the other site, and attributed this difference to geometric characteristics and sight distance. However, there is no detailed relationship provided between geometric characteristics and capacity estimated.

Kerner (2004) stated that capacity at a bottleneck was determined by the transition from free flow to synchronized flow and is random by nature. The probability of transition depends on the traffic demand and on the time interval of the observations, thus, there was an infinite number of freeway capacities in free flow at bottlenecks. The probability of transition from free flow to synchronized flow and the corresponding capacity at a bottleneck for a specific site are illustrated in Figure 2-2.

Corresponding to Figure 2-2, the capacity of the freeway increased with increasing breakdown probability: the minimum capacity happens when the breakdown probability is 0 and the maximum capacity happens when the breakdown probability is 1. He also stated that along a homogeneous road the freeway capacity depends on which phase the traffic is in and thus defined maximum freeway capacities in free flow, in synchronized flow (all vehicles in the same direction have the same time-independent speed) and in wide moving jams (congestion propagates upstream through bottleneck with a characteristic speed of about 16 km/h). The freeway capacity in free flow phase was the flow at which the transition from free flow to synchronized flow occurs; The capacity in synchronized flow was the maximum possible flow rate; The capacity at wide moving jams was equal to the flow rate in the wide moving jam outflow. In conclusion, the author defined capacity at different locations and traffic states and treated it as a
variable, similar to that proposed by Lorenz and Elefteriadou (2001), Persaud and Hurdle (1991). However, the reasons of the definitions were not fully explained.

Brilon (2005) concluded from extensive data analysis that capacity is a variable, even under identical environmental conditions, and considered only the volumes that cause the breakdown as capacities. This implies that capacity may not be the highest volume available. Similarly, Brilon et al. (2005) proposed that capacity should be understood as the traffic volume below which traffic still flows and above which the flow breaks down into stop-and-go or even completely stop traffic. They estimated capacity using the Product Limit Method based on the statistics of lifetime data analysis and found that the capacity of a freeway section is Weibull-distributed, however, different freeways might have different parameters. The estimated capacity distribution is shown in Figure 2-3. It can be seen from Figure 2-3 that capacity of the freeway increases with increasing breakdown probability, similar to the conclusion of Kerner (2004).

Corresponding to a certain breakdown probability, the capacity of the section with variable speed limit is much higher than that without speed limit.

Banks (2006) used average time gaps (average time separations between the rear of a vehicle and the front of a vehicle following it), average passage times (average times for vehicles to pass a point), and lane flow distributions as variables to link geometric, vehicle population, and driver population characteristics to capacity flows. Two types of “capacity condition” – pre-queue flow and queue discharge flow – were considered. One major finding was that average time gaps appeared to be the most important flow characteristic for explaining variations in pre-queue flow and queue discharge flow, and average passage times, on the other hand, were not correlated with
flows. Finally, there was a strong, near-linear negative relationship between critical lane flows and critical lane average time gaps in both pre-queue and queue discharge flow.

Yeon et al. (2007) proposed four parameters to define capacity, which were maximum pre-breakdown flow, breakdown flow, maximum queue discharge flow and average queue discharge flow. They concluded from statistic analysis of speed and volume data that the average capacity flows during the AM peak period were slightly higher than that during the PM peak period, and that capacity flows for both peak periods were higher than those during non-peak periods. Among the four parameters of capacity, the average maximum pre-breakdown flows at all the locations were the highest, while the average queue discharge flows were the lowest.

Literature related to the definition of capacity in other research is summarized in Table 2-1.

In summary, field data had shown that there was variability in the maximum flow observed, in the range of several hundred vehicles per hour per lane. Four different flow parameters (maximum pre-breakdown flow, breakdown flow, maximum discharge flow and average queue discharge flow) were used to define the capacity of a freeway facility. The parameters might vary widely even at the same site, probably due to the microscopic characteristics of traffic and drivers, such as individual spacing, time headway, speed of individual vehicles in traffic stream and their variability. However, there is limited research on a quantitative explanation of the flow variance and on the impact of site characteristics on the maximum flow distribution form.

Measurement of Capacity

Critical issues for measuring capacity include the measurement location, the measurement time and the presence or absence of queues both upstream and
downstream of the measurement location (Tian et al., 2005). This subsection reviews literature related to the time, location, and method of capacity measurement, as well as the “two capacity” hypothesis.

Many researchers studied on the time and location of capacity measurement. Hall and Agyemang-Duah (1991) stated that the queue-discharge capacity must be measured at an active bottleneck location, which was characterized by the presence of an upstream queue and the absence of a downstream queue, and that free-flow capacity measurement should be restricted to the period prior to breakdown when demand was approaching a maximum. For bottlenecks related to merge sections, the flow measurement location should be somewhere downstream of the freeway merge, so that both freeway and ramp flows were counted (Tian et al., 2005).

Ringert and Urbanik Π. (1992) developed an empirical model for estimating the maximum sustainable flow at freeway bottlenecks in Texas, based on the analysis of data collected at several sites. Several results had been obtained. Firstly, freeway bottlenecks were the best location for measurement of capacity, as a result of their ability to determine the transition to queue discharge flow. Secondly, turbulence caused by imbalance of traffic (e.g., merging and weaving activities) might prematurely transitions the flow from free flow into queue discharge flow. Thirdly, queue discharge flow was the best estimate of maximum sustainable flow, as it had significantly lower variability than free flow rate, and thus was recommended as the capacity. And finally, much lower flow rates might occur if the study site was affected by downstream congestion. Based on the analysis, a speed-flow model was developed and the maximum sustainable flow rate was determined to be 2,200 pcphpl. However, the
results were based on several assumptions, such as no downstream congestion and that traffic conditions remaining uncongested.

With regard to the method of capacity measurement, Hyde and Wright (1986) proposed to estimate capacity by simply selecting the maximum flow rate measured over the observation period or applying expected extreme value method. However, the capacity estimated by this method highly depends on the duration of the average interval. Van Arem and Van der Vlist (1993) proposed to estimate the current capacity by updating the fundamental diagram, which was determined under predefined conditions. By calibration, this method can reflect the variance in capacity under different traffic compositions and weather conditions. But this method was aimed at estimating capacity while the traffic was free-flowing, and thus can not estimate the real-time capacity during congested conditions.

Minderhoud et al. (1997) summarized many methods to estimate capacity (direct empirical and indirect empirical), and found the product limit method, empirical description method and fundamental diagram method to be the ones with highest validity. Further, they discussed the application of product limit method, in which different traffic states needed to be observed and the bottleneck location should be chosen, in estimating capacity and its distribution based on both flow and speed data. From this viewpoint, this method can provide a good estimation of the capacity because it utilizes the traffic state information and gives a distribution of capacity instead of a single value.

The above methods can estimate theoretical capacity, but are not adequate for real time applications such as ramp control, thus, real time capacity estimation method
is required. Liu et al. (2008) presented a real time freeway estimation method in freeway ramp control. When the freeway section was uncongested with occupancy at the present less than the critical occupancy (averaged around 15.5), 95% percentile of theoretical capacity, which was the maximum flow measured from the flow-occupancy diagram, was defined as the capacity. When congestion begins with occupancy exceeding the critical occupancy, operational capacity, which was the actual maximum flow rate and obtained through moving average method based on real time measurements, was defined as the capacity. They tested the methodology on the Stratified Zone Metering (SZM) strategy through micro-simulator and found that the control strategy improved by decreasing delay and increasing speed.

A number of investigations had proven the existence of different capacities under uncongested and congested traffic conditions (i.e. the “two capacity” hypothesis). The implication of two capacities affects the definition and value of capacity. Banks (1990) supported the hypothesis that maximum flow rates decreased when queues form, and analyzed the “capacity drop” phenomenon for different North-American freeways. Capacity drop values of between 3 and 6% were measured. Hall and Agyemang-Duah (1991) investigated the issues of whether there was maximum flow reduction when a queue forms and where capacity could be properly measured. They concluded from data analysis that there was about a 5 to 6 percent reduction of maximum flows at bottlenecks after the onset of congestion, and indicated that capacity can only be measured in a bottleneck but not in a queue. Ponzlet (1996) analyzed traffic flow on German freeways to test whether the “capacity drop” phenomenon existed, and a 6% drop of flow rates for 5-minute was determined. Brilon and Zurlinden (2003) analyzed
the capacity drop by comparing the stochastic capacity to flow rates during congested flow, and computed an average of 24% capacity drop, which was very high compared to other researchers’ results. Regler (2004) analyzed the capacity drop value using a distribution of breakdown flows and queue discharge flows, and found an average drop of 250 veh/h in 5-minute flow rates.

Later, Zhang and Levinson (2004) examined the two-capacity hypothesis about flow drops after breakdown with a dataset of twenty-seven bottlenecks. The results showed that the percentage flow drops at various bottlenecks follow a normal distribution with mean 5.5% and standard deviation 2.3%. Brilon et al. (2005) reported that the capacity drop was stochastic, and capacity was Weibull-distributed with a nearly constant shape parameter. They further extended this stochastic concept into reliabilities of freeway networks by stating that a freeway may operate at the highest expected efficiency if it was only loaded to 90% of the conventionally estimated capacity.

**Summary**

In summary, several research studies the definition of freeway capacity and its distribution. The HCM is the first and most often used document to quantify the concept of capacity. The definition of capacity in this manual varies overtime due to the improvement of facilities as well as the realization of the variance of capacity under different conditions. With respect to the inadequacy of the capacity definition in the HCM, many researchers suggested defining capacity under different traffic conditions, however, no consensus have been achieved. Four different flow parameters (maximum pre-breakdown flow, breakdown flow, maximum discharge flow and average discharge
flow) are most often used to define capacity, and microscopic characteristics of traffic and drivers are considered to be the reasons of capacity variance at the same site.

The methods used in measuring capacity include observing maximum flow rates, updating the fundamental diagram, applying product limit method and using real time two-phase capacity estimation method, etc. Freeway bottlenecks are considered to be the best location for measurement of capacity.

However, there are still some problems not resolved by previous research:

1. Some research indicated that microscopic characteristics of traffic and drivers, and average time gaps to be the main reason in capacity variance. More research which quantitatively analyzes the link between the flow variance and its reasons is necessary.

2. Given the concept of freeway capacity randomness, it is necessary to know more about the form of the capacity distribution function. Most of the research does not examine the form of capacity distribution, and there is no research on the impact of site characteristics on the form.

The Process of Breakdown and Breakdown Models

As referred earlier, freeway capacity is also related to flow breakdown, which is the transition between uncongested and congested flow conditions, and some researchers suggested taking into consideration the breakdown probability in defining capacity (Lorenz and Elefteriadou, 2001). The relationship between the probability of breakdown at freeway bottlenecks and flow rate is of interest to both capacity analysis and freeway operation. This section reviews previous research related to the process and probability of breakdown.

Banks (1991) conducted a study of four bottlenecks in San Diego and suggested that breakdown was triggered by slow-moving vehicles that cause unstable speeds in dense platoons. He also observed that breakdown occurs upstream of the merge for three cases and both upstream and downstream of the merge for one case.
Elefteriadou et al. (1995) examined the breakdown occurrence at two freeway-ramp merges, using data collected from four video cameras along the ramp merge area. Investigation of the data revealed that the breakdown occurrence was associated with the presence of vehicle clusters coming from the on-ramp. The ramp vehicle cluster would “force” its way into the freeway and result in lower speeds. Based on this mechanism, they proposed a probabilistic model, which was a smooth S-curve and a function of freeway and ramp flow rates.

Kerner and Rehborn (1997) studied data collected from detector sensors in the vicinity of freeway ramp merge sections on German highways. They stated that in the vicinity of bottlenecks, the breakdown occurs due to local speed decrease and density increase that is observed when on-ramp vehicles squeeze on the highway, or due to unexpected speed decrease and lane changing activity. They stated that the latter can result in breakdown even away from a bottleneck.

Persaud et al. (1998, 2001) examined the breakdown process over a large number of rush hours at three freeway bottlenecks using detector data. They observed that breakdown was always associated with a sharp decrease in speed and flow and the formation of vehicle queues upstream of the bottleneck.

Chung and Cassidy (2004) reported that, based on a study of data from nine days, the breakdown was triggered by high vehicle density that begins from an individual lane and spreads across the freeway segment as drivers change lanes to avoid decelerating.

Several models have been developed to predict the occurrence of breakdown. Elefteriadou et al. (1995) developed a model which assumed that breakdown occurs if at least one vehicle on the freeway is forced to reduce its speed by 16 km/h or more.
The probability of breakdown was computed as a function of the cluster size of on-ramp vehicles and the freeway flow through three steps: calculating the probability of occurrence of all possible cluster size (3 to 15 vehicles), calculating the probability that at least one vehicle is present at the critical area of the freeway (the shoulder lane), and estimating the drivers’ possible actions that they may take as the cluster of vehicles approaches the freeway from the ramps. The resulting model provided the probability of breakdown as a function of freeway and ramp flow rates, and had a smooth S-curve shape. The model is illustrated in Figure 2-4.

Evans et al. (2001) used Markov chains to develop the probability distribution of breakdown at ramp-freeway junctions based on zonal merging probabilities with respect to the vehicles traveling on the throughway.

Lorenz and Elefteriadou (2001) conducted an extensive analysis of speed and flow data at two freeway sections in Toronto, Canada. They observed that the breakdown occurred at various levels of demand and the probability of breakdown increased with increasing flow rate, as shown in Figure 2-5.

Brilon et al. (2005) investigated the relationship between the probability of breakdown and flow using the Product Limit Method (PLM) and developed models that predict the probability of breakdown as a function of flow.

Elefteriadou et al. (2009) developed breakdown probability models for five freeway sections in North America using the PLM, and incorporated the models in ramp metering algorithms. The probability of breakdown was developed as a function of the ramp demand and the freeway demand (considered individually), or as a function of the sum of the ramp and freeway demands upstream of the critical ramp junction, or as a
function of the occupancy upstream of the critical ramp junction. The authors suggested that the type of breakdown probability model to be developed depends on the application purposes. For example, they recommend building a volume-based breakdown model if the breakdown probability model is to be incorporated within a ramp metering algorithm that uses freeway volume for selecting the metering rates; they recommend building an occupancy-based model if the ramp metering algorithm to be applied is based on occupancy.

In summary, previous research has associated breakdown with driver actions (braking, lane changing). Breakdown can be observed through a decrease in speed and flow, an increase in vehicle density, and the formation of vehicle queues upstream of the bottleneck. There have been several models reported that predict the occurrence of breakdown as a function of flow, including most recently the PLM. These models were developed based on several different parameters, including flow and occupancy. The literature review search did not reveal any applications of the PLM for predicting incident-induced breakdowns.

**Freeway Capacity and Operation under Incident Conditions**

Incidents that block freeway lanes or create other impedances to driving behaviors can create significant negative impact on operational conditions, such as increased queues, delays and pollution. Incidents can also reduce freeway capacity by blocking one or more lanes and/or the adjacent shoulder lane. This section reviews literature related to the impact of incidents on freeway capacity and the relationship between incidents and operations, and proposes problems that deserve further research.
Impact of Incidents on Freeway Capacity

When an incident occurs, the capacity at the road section may decrease, as probably some lanes are closed. Even if no lane was closed (incidents occur on shoulder), vehicles from upstream may slow down near the incident location, causing capacity reduction as a result of speed reduction. Estimating the freeway capacity following an incident is important for effective traffic management. This subsection reviews the impact of incidents on freeway capacity.

The HCM 2010 (also the HCM 2000) addressed the issue of capacity reduction due to incidents as presented in Table 2-2. This manual reported that the effect of an incident on capacity depends on the length of blocked road section, the number of lanes, drivers slowing down, and the rubbernecking factor.

Goolsby (1971) analyzed 27 incidents that occurred between 1968 and 1969 on a 6.5-mile section of the Gulf Freeway in Houston using 1-minute volume counts. Based on these data, Goolsby estimated that an incident or disabled vehicle blocking one of three lanes will result in an average capacity reduction of 50%, an incident blocking two lanes of three lanes will reduce capacity by an average of 79%, and an incident or disabled vehicle blocking the shoulder lane on a three-lane segment was found to reduce capacity by an average of 33%. The results obtained by Goolsby are consistent with that given by the HCM, and are still widely used by transportation professionals. Both the HCM and Goolsby’s research report incident capacity reduction as a deterministic value.

Chin et al. (2002) estimated the number of closed lanes based on the type of crash as well as the number and type of vehicles involved. They assumed that a fatal or injury crash involving more than one vehicle always result in lane closures. The
probabilities of lane closure and number of lanes closed are illustrated in Tables 2-3 and 2-4. However, the estimated probabilities are used for macroscopic estimation of delay and are not suitable for microscopic analysis such as breakdown phenomenon at a specific freeway site.

Smith et al. (2003) measured capacity reduction caused by over 200 incidents that occurred on urban freeways in the Hampton Roads region of Virginia. The authors defined the minimum 10-minute flow rate (by moving average of five successive flows to reduce variation) measured in the bottleneck created by an incident as incident capacity and found that: an incident blocking one of three freeway lanes resulted in a mean capacity reduction of 63% (significantly larger than the widely accepted value of 50% as proposed by the HCM and Goolsby’s research in the 1970s), while an incident blocking two of three freeway lanes resulted in a mean capacity reduction of 77% (similar to Goolsby’s research and smaller than that given by the HCM). They also stated that it was more appropriate to model incident capacity reduction as a random variable and not a deterministic value due to the variation of traffic flow, and suggested that the beta distribution provides a good representation of incident capacity reduction for one or two of three lanes blocked.

Chin et al. (2004) estimated capacity loss on freeways and principal arterials that resulted from fatal and non-fatal crashes, vehicle breakdowns, and adverse weather, etc. They assumed that a fatal or injury crash involving more than one vehicle always results in lane closures (i.e., probability of lane closures = 100 percent for injury crashes). Then they estimated the number of lanes closed according to the location of crash and crash severity, and calculated capacity remaining after incidents based on
the total number of lanes available and the number of lanes affected by the crash, as shown in Table 2-5. In summary, the authors studied capacity loss caused by incidents based on some assumptions about the number of lanes closed. The study does not consider geometric conditions in capacity estimation.

Potter et al. (2007) analyzed the number of lanes affected by different incident types. The relationship between incident types and number of lanes affected is shown in Figure 2-6. Figure 2-6 A) shows the percentage of different incidents types in 2005. It was observed that over half (56%) of the incidents were stalls. The next most common types of incidents were crashes (17%) and debris (13%). Figure 2-6 B) shows the number of lanes broken down by incidents for the year 2005. It was observed that 62% of incidents did not block any lane. This observation is likely related to the high occurrence of stalls as shown in Figure 2-6 A): 33% of incidents blocked one lane, and only 5% of incidents blocked two or more lanes.

Hadi et al. (2007) examined the simulation modeling of incidents and their impacts on capacity. Three types of simulation models were investigated: CORSIM, AIMSUN and VISSIM. These allow the users to simulate incident blockages either explicitly or by using other events that have similar effects on traffic operation. The results showed that all the three models need to calibrate parameters for the incident location to produce the capacity reduction similar to that estimated by the HCM 2000 or field studies. Only CORSIM included incident calibration parameters (rubbernecking factor). AIMSUN and VISSIM did not have incident calibration parameters but had other modeling capabilities (the speeds of the unblocked lanes at the incident locations) that can be used to calibrate the models to produce the identified reduction in capacity due to incidents.
Knoop et al. (2008) collected data about two incidents by helicopter to analyze the impact of incidents on freeway capacity. The results showed that the capacity (outflow) at the opposite direction of the incident location was reduced by half by the “rubbernecking” effect (drivers slow down to see what is happening). The capacity of the road section in the direction of the incident was reduced by more than half, as not all lanes are in use. While most researchers studied the capacity reduction only in the direction of the incidents, this research showed that the effects in the opposing direction could be of the same order. However, this effect of incidents on the opposing direction is to be examined on other freeways. This paper does not make a comparison of the incident capacity reduction observed to the values in the HCM 2000.

**Relationship between Incidents and Freeway Operations**

Freeway operational condition can impact safety. Al-Deek et al (1995) investigated the relationship between freeway geometry and the location, type, and time period of incidents. They found from statistical analysis that the upgrade freeway segments had significantly higher incident rates than level or downgrade segments, off-ramps had significantly higher incident rate than on-ramps or no ramps, and peak periods had significantly higher incident rates than off-peak periods.

Many previous researches suggested that there was a positive association between speed and the risk of crash involvement. For example, Solomon (1964) studied the relationship between crash and speed. They found that crash rates took a U-shaped form against the difference in speed and average speed. Crash rate was maximum for vehicles with speeds of more than 35 mph below the average and minimum for speeds of 5 to 10 mph above the average. Kloeden et al. (1997) concluded that "In a 60 km/h
speed limit area, the risk of involvement in a casualty crash doubles with each 5 km/h increase in traveling speed above 60 km/h”.

There is also extensive research on the relationship between incident frequency and operational conditions. Chang et al. (2000) clarified the relationship between volume to capacity ratios (v/c) and accident rates at various freeway facility sections, based on data on Shingal-Ansan freeway in Korea from 1992 to 1997. The results showed that the relationship between accident rates and v/c ratios represented a U-shaped pattern for all sections (basic freeway, tunnel, and toll gate sections), as illustrated in Figure 2-7. They further observed that the accident rate of the toll gate section was generally higher than that of other sections, and that there was no significant difference in accident rates between the basic freeway and tunnel sections when the v/c is between 0.5 and 0.8. Moreover, the basic freeway, tunnel, and toll gate sections had the minimum accident rates when the v/c rations are 0.78, 0.75, and 0.57, respectively.

Garber and Wu (2001) studied the relationship between the crash probability and traffic as well as geometric characteristics based on a selected section of Interstate 64 within Norfolk, Virginia Beach, and Chesapeake in Virginia. The study found that traffic volume, occupancy, standard deviation of speed, and exposure had significant positive relationships with the number of crashes, while speed had significant negative relationship with the mean of number of crashes. They also found that the stochastic regression modeling methods could be used to describe the probabilities of crash events. Moreover, the authors stressed the use of corresponding traffic data (those occurring at the time of the crashes), as it reflected the true influences of independent
variables on the occurrence of crashes. These findings may help to determine the criteria in incident detection.

Golob et al. (2002) developed a tool, called FITS (Flow Impacts on Traffic Safety), to evaluate the effect of changes in traffic flow on traffic safety. They classified traffic flow into several conditions, and got the probabilities of different kinds of crashes under different conditions. Continuing their study, Golob et al. (2003) used the tool FITS to forecast the types of crashes that are most likely to occur for the flow conditions being monitored. In applying the tool, they found that crash rates and types were related to traffic flow patterns: the highest crash rate happened under heavily congested flow (5.99), then variable-volume congested flow (3.21), and then variable-speed congested flow (2.97). The lowest crash rate happened when flow was approaching capacity (0.55). They also studied the relationship between crash types and mean flow as well as median speed. For example, at flows higher than mean flow and speed higher than median speed conditions, 78.8% of the crashes happened might be two & multi-vehicle rear-end & lane-change crashes, 57.8% of the crashes happened might be two-vehicle lane-change & rear end crashes. The proposed tool developed detailed relationships among crash rates, crash types and traffic flow patterns, and can aid in identification treatments that aimed at enhancing safety. However, future research is needed to compare the present method with other approaches that have been recently developed. Another limitation is that the FITS tool applies only to urban freeways with at least three lanes in each direction.

Golob and Recker (2003) used linear and nonlinear multivariate statistical analyses to determine how the incident types were related to traffic flow, weather and
ambient lighting conditions, based on data from heavily used freeways in Southern California. Results showed that the types of collisions were strongly related to median traffic speed and to temporal variations in speed in the left and interior lanes. Controlling for weather and lighting conditions, there was evidence that incident severity was influenced more by volume than by speed.

Incidents also have impacts on freeway operations. Menendez and Daganzo (2004) used the kinematic wave theory and some dimensional analysis to evaluate how the location and duration of an incident influence delays near a recurrent bottleneck. The incidents are classified according to the delay they incurred to the system. The delay was modeled by the intrinsic characteristics of an incident such as capacity at the incident sites, length of rubbernecking zone created by the incident and free flow speed at the incident site, together with the incident location and duration. However, how to get the capacity at incident site and the length of rubbernecking zone were not illustrated.

Research on the relationship between congestion and safety is rich (Jernigan, 1998): Tedesco et al. (1993) noted that reducing congestion can lead to a reduction in secondary crashes. In the event of a crash, the risk of a secondary crash was increased by more than 600%. Persaud et al. (1998) stated that preserving un-congested operations on freeway resulted in both improved safety and travel time benefits to the travelers. On the other hand, when breakdown occurred and flow transited to congested operation, travel times increased as queues form, and incidents were more frequent. At two freeway corridors, one in Los Angeles and the other in the Bay Area, incident-related delay was found to be 13% to 30% of the total congestion delay experienced during peak periods (Prevedouros et al., 2008).
Summary

Incidents may lead to capacity reduction, as they usually block lanes and affect other drivers. There is one paper defining the minimum 10-minute flow rate (by moving average of five successive flows to reduce variation) measured in the bottleneck created by an accident as incident capacity. Previous research estimated capacity remaining after an incident based on the number of lanes closed by the incident, and reported incident capacity reduction as a deterministic value. There is also suggestion to model incident capacity reduction as a random variable and suggests the beta distribution for one or two of three lanes blocked. The capacity remaining after incidents under various conditions is summarized in Table 2-6.

There has also been research predicting the probability of lane closure and number of lanes closed from type of crashes and number of vehicles involved. The results might be used in this dissertation for sites without information of the number of lanes closed or affected.

On the one hand, freeway operations have impact on incidents. On the other hand, incidents also have impact on freeway operations. Incident rates and types are related to operational conditions such as volume, v/c ratio, speed variation and congestion condition. Congestion can be modeled by the characteristics of incidents. Most of the studies analyze the whole freeway instead of different road sections (e.g., basic section, weaving section on freeway). Limited research studied the effect of incidents on traffic flow under congested conditions, for instance, the impact of incidents on flow breakdown. Future research on the impact of incidents on traffic flow at freeway bottlenecks under congested conditions is necessary.
Further research is necessary in the following areas:

(1) Incidents on freeways occur often and they have substantive impact on freeway capacity. However, there is limited previous research on defining freeway capacity under incident conditions. Thus a definition of freeway capacity, which considers the traffic flow patterns, geometric characteristics and incident conditions, is to be proposed.

(2) Apart from the number of lanes affected by incidents, it is necessary to estimate capacity remaining after an incident based on other factors, such as geometric conditions, the incident duration and type, etc, as there is no clear relationship between them.

(3) It is desirable to model capacity remaining after an incident as a random variable rather than a deterministic value because of the variations associated with the incident such as traffic control, local conditions and driver behaviors.

(4) For sites without information of the number of lanes closed/affected in the dissertation, investigate the potential of predicting the number of lanes closed from incident type and number of vehicles involved.

(5) Investigate the relationship between incident frequency and traffic operations at freeway bottlenecks. For example, the frequency of incidents at bottleneck and non-bottleneck locations.

Identification and Verification of Incidents

The identification and verification of incidents are important in traffic/incident management practice. This section reviews literature related to the detection, prediction, and verification of incidents.

Incident Detection

Incident detection is the process of determining the presence and location of an incident. The idea has been around since the mid-1960s and early 1970s as part of standard traffic/incident management practice. Incident detection algorithms are classified into different types, including comparative, statistical, time series, theoretical models, and artificial intelligent algorithms, etc. (Mahmassani et al., 1999, Ozbay and Kachroo, 1999).
Comparative or pattern recognition-based algorithm. These were first developed by traffic engineers and are by far the more common types of algorithms. Their basic principle is that an incident will create increased occupancy upstream and decreased occupancy downstream of the incident. California algorithms are the best-known comparative algorithms. The basic California algorithm was originally developed in the late 1960s (Payne et al., 1976). It detected incidents by comparing the measured occupancy from two adjacent detectors. The general logic of this algorithm is depicted in Figure 2-8.

In Figure 2-8, T1, T2 and T3 are site-specific occupancy values that need to be calibrated from empirical data. Payne and Tignor (1978) modified the California algorithm and published 10 new versions, among which No. 7 and No. 8 were the two that performed best. Mahmassani et al. (1999) reviewed the literature and concluded that the comparative algorithms work best under moderate-to-heavy traffic conditions, but cannot handle fluctuating traffic demands efficiently, as they rely on static thresholds. These algorithms are most widely implemented and have been installed in California, Chicago, and Texas (Martin et al., 2001).

Statistic algorithms. These algorithms model the stochastic traffic flow patterns obtained from loop detector data. Such algorithms include the Standard Normal Deviation algorithm that is based on mean and standard deviation of occupancy (Dudek et al., 1974), and the Bayesian algorithm that is based on the relative difference in occupancy (Levin and Krause, 1978).

Time series algorithms. This kind of algorithm considers the recent history of a traffic variable and employs statistical forecasting of traffic behavior to provide short-
term traffic forecasts (Mahmassani et al., 1999). Such algorithms include the High Occupancy (HIOCC) algorithm that is based on occupancy (Collins et al., 1979), the Autoregressive Integrated Moving Average (ARIMA) algorithm that uses occupancy (can also use volume or speed) to detect incidents (Ahmed and Cook, 1982), and the Double Exponential algorithm that uses one-minute average volume, occupancy and speed (Cook and Cleveland, 1974). The Double Exponential algorithm has been implemented in Toronto (Martin et al., 2001). Another is the Low-pass Filter algorithm developed by Stephanedes and Chassiakos (1993), which is based on occupancy.

**Traffic (modeling) and theoretical algorithms.** These employ the basic theories of traffic flow characteristics. The most notable one among these is the McMaster algorithm (Persaud and Hall, 1989). It was based on the catastrophe theory that sudden changes could occur in one variable of interest (in this case the speed) while other related variables (in this case the flow and occupancy) exhibit smooth and continuous change. Generally, this algorithm classifies traffic operation on the freeway into four states using volume and occupancy. Once congested operation is detected at any detector, operations at the downstream detectors are evaluated to identify whether the congestion was caused by an incident or not. The algorithm requires calibration of the boundaries separating different traffic conditions at each detector (Mahmassani et al., 1999, Parkany and Xie, 2005). It has been implemented in Minnesota (Martin et al., 2001).

**Other algorithms.** Artificial intelligent based algorithms use input data of traffic flow variables to get associated output results. This kind of algorithm includes Neural Networks (Dia and Rose, 1997) and Fuzzy algorithms (Chang et al., 1994). Wavelet
techniques, which are usually used in detecting changes in signals in electrical engineering, have also been used to detect incidents (Teng and Qi, 2003, Jeong et al., 2009). Mahmassani et al. (1999) suggested combining existing algorithms to achieve the advantages of different algorithms in incident detection. This process is termed Algorithm Fusion.

Automatic incident detection algorithms are usually evaluated using three measures: detection rates (DR), false alarm rates (FAR) and mean time to detect (MTTD). DR is the proportion of detected incidents to the total number of actual incidents during a given time period. False alarm rate (FAR) has different definitions for different applications. It has mostly been defined as the ratio of the number of false alarms to the total number of algorithm applications (off-line FAR). It has also been defined as the fraction of the number of false alarms to the total number of declared incident alarms, including all correct and false alarms (on-line FAR). Finally it has also been defined as the number of false alarms per day or per hour. MTTD is defined as the average time spent from the moment an incident occurs until the time that the algorithm declares this incident, averaged for all incidents detected over a period of time and measured in minutes (Balke, 1993; Parkany and Xie, 2005). In previous research, the algorithms are usually tested with simulation data or/and field data. The best performing commonly used algorithms are summarized in Table 2-7. This table uses the definition of off-line FAR. However, it is not clear based on the literature reviewed which of the algorithms are tested with simulation data or field data.

As shown in Table 2-7, almost all the algorithms have the limitation of false incident prediction. This occurs because there was a large overlapping area in which
traffic stream variable measures are the same during both non-incident and incident operations (Cook and Cleveland, 1974). Guin et al. (2004) conducted a survey to personnel in Transportation Management Centers all over the United States and the Ontario Ministry of Transportation at Ontario, Canada, to evaluate the status of existing implementation of incident detection algorithms. There were 32 centers responding to the survey, with the viewpoint that the unacceptably high false alarms rate is the major deterrent of incident detection algorithms, and, on an average, a maximum of three false alarms per hour and an average of ten false alarms per day are considered acceptable. Brydia et al. (2005) also conducted an automatic incident detection survey. With respect to the acceptable number of false alarms that operators are willing to accept, per day, per station, the predominant (60%) answer was two to five.

In conclusion, there is a significant amount of research on incident detection. The California algorithms and McMaster algorithms are the most widely known and relatively simple compared to more recently-developed algorithms. These two are often used as a standard to evaluate newer algorithms. The literature review search did not reveal any applications of the PLM for the detection of incidents. It is also concluded from the literature that occupancy is most often used in incident detection. Other traffic parameters used for incident detection include volume, speed, and standard deviation of occupancy.

**Incident Prediction**

Real-time incident prediction is the process of using current operational data analysis to predict the probability of an incident during the following time period. Several methods have been developed and used to predict incident occurrence.
Liu (1997) predicted the likelihoods of two types of incidents for a particular period on a specific section on the freeway using a binary logit model. Parameters found to be significant in predicting overheating vehicle incidents included peak factor, merge, temperature, rain, and speed variance, the $\rho^2$ value is 0.215. Parameters found to be significant in predicting vehicle crash incidents included merge operations, visibility, and rain, the $\rho^2$ value is 0.14. Traffic variables such as flow are not significant in crash incident prediction. However, the data are hourly recorded and thus lack precision. As pointed out by the author, the hourly data levels out the large changes in other parameters such as average speed, flow and speed variance at the beginning of incident. The developed incident likelihood prediction models are evaluated by both mathematical formulation and simulation, and found that the reduction in incident waiting time (incident detection plus response time) is significant when the models are used. The model results are not compared with other researches.

Oh et al. (2001) used a non-parametric Bayesian modeling approach to predict the real-time incident likelihood based on real-time traffic data and incident data for the I-880 freeway in Hayward, California. The authors classified traffic conditions into two patterns: disruptive (a 5-minute period right before an incident) and normal (a 5-minute period 30 minutes before an incident occurrence). They used the 5-minute standard deviation of speed (selected from six candidate precursors: the mean and standard deviation of three traffic parameters flow, occupancy, and speed) to identify incidents. The model is applied to real-time detector data, and the results show that there are some false predicted incidents. The model results are not compared to other researches. However, insufficient data led to some key assumptions and thus limited
the scope of the study. Moreover, it used only the standard deviation of speed as the incident indicator, which cannot completely define traffic operation.

Lee et al. (2003) developed a log-linear model using categorical variables for incident prediction. They investigated three crash precursors in the model: the average variation of speed on each lane (measured by the coefficient of variation of speed, CVS), traffic density (D), and average speed difference between upstream and downstream ends of the road section (Q). Each parameter was classified into several categories. They also investigated exposure (the product of daily traffic volume and the length of each road section), and two control factors: road geometry and time of day. They found the average variation of speed on each lane (CVS) and traffic densities to be significant incident predictors, and that CVS has a relatively longer-term effect on crash potential than either D or Q. Also, higher-level crash precursors contribute to higher crash potential than lower-level crash precursors. The developed model was calibrated for historical crash data using the maximum likelihood estimation method, but not used to predict crash potential in real-time using the current traffic flow data. Another problem is that, instead of predicting the potential of an incident occurring during the following time intervals, the model predicted the number of incidents per vehicle-kilometers of travel.

Balke et al. (2005) used a binary logit model to predict incidents using loop and weather data. They tested average volume, average occupancy, average speed, and CVS, using different detection time and moving average window size. The estimation results indicated that occupancy and average variation of speed are potential precursors of crashes, and the model fitted best when using 15-minute detection time
and 5-minute moving average. The modeling results are shown in Table 2-8. The model was tested with both incident and non-incident data. An example of a five-day data set test is shown in Figure 2-9. A collision incident was reported on October 12, 2004, at 1:06 PM. The authors concluded that the model predicts high likelihoods of collision quite accurately. However, how to distinguish incidents from congestion was not investigated. This model might cause false prediction of incidents during unstable traffic flow, which may have high average CVS value. The model results are not compared with other researches.

Pande et al. (2005) used logistical regression model to measure the hazard ratio at different locations around the incident and at different time slices 30-minutes prior the incident, based on 2046 crashes collected from 4 years on Interstate 4 in Orlando, Florida. The log CVS, standard deviation of volume and average occupancy were found to be the most significant variables affecting the crash odds. The parameters calculated from 5-minute level were more significantly correlated with crash than that calculated from 3-minute level. It was observed from the spatiotemporal analysis that the crash odd increased as the time and location of the crash were approached. Thus the authors recommended to implementing measures to reduce the speed variance, once a potential crash location was identified in real time. Abdel-Aty et al. (2005) extended this methodology to predict freeway crashes under low-speed regime (near 25 mph) and high-speed regime (near 55 mph). Different models were developed for the two regimes using logistic regression and it was found that different parameters were kept in the final models.
Kuchangi (2006) used a categorical log-linear model to predict the incident probability. The parameters used were exposure in vehicle-miles of travel, CVS, occupancy, geometric indicator (on-ramp, off-ramp), peak hour factors. Each parameter had several categories. The results showed that the rate of crash occurrence increased with CVS and occupancy. Occupancy was regarded as a stronger precursor to predict incidents than CVS, as it had larger range between maximum and minimum estimates. Crash rate was lower during non-peak hours than during peak hours, and was lower at straight sections or sections without on-ramps and off-ramps than at curved sections or sections near to ramps. Overall, roadway-type indicator and peak-hour indicator were less sensitive when compared to CVS and occupancy. The parameter ‘exposure’ was found to be statistically insignificant. However, the developed model predicted the number of crashes at different categories of the parameters but not the real-time crash likelihood, and the model was not validated.

Pande et al. (2011) examined the relationship between crash risk and real-time traffic variables from a freeway corridor based on loop data collected from I-4 eastbound in Orlando, FL. They considered nine parameters (average, standard deviation and logarithm of coefficient of variation for speed, volume and occupancy), and developed a variable selection procedure based on random forests. Average occupancy of upstream station, average speed and coefficient of variation of volume for downstream stations were found to have significant effect on crash risk. The developed model is also applied to three other freeway corridors (I-4 westbound, and I-95 north and southbound). The results showed that the model developed for I-4 eastbound corridor works good for the I-4 westbound corridor, however, the performance is not as good for I-95. They then
suggested that the same model for crash risk identification may only work for corridors with very similar traffic patterns.

There had also been research regarding incident type prediction. Pande and Abdel-Aty (2006) used neural networks to develop models for predicting the rear-end crash risk. They found that, in congested conditions, the average occupancy and average variation of speed at the location (loop detector station) of interest were the most important variables affecting the rear-end crash risk. While in non-congested conditions, the speed difference was very important, as the crash risk was increased when faster moving vehicles approaching slower moving vehicles. Thus average speeds at the location of interest and both upstream and downstream of this location were required. They also used a single neural network model to predict lane-change crashes, and found that the main factors affecting the lane-change crash risk were the average speeds upstream and downstream of the station of interest as well as the difference in the lane occupancies across each individual lane on the freeway. The higher the absolute difference in the lane occupancy across adjacent lanes on freeway, the higher the chance of having a lane-change related crash.

However, there was one research by Luo and Garber (2006) reporting that incidents cannot be predicted. They used three pattern recognition methods: the K-means clustering Method, the Naïve-Bayes Method and the Discriminate Analysis, to differentiate crash-leading pattern from normal non-crash pattern. They extracted 391 crashes as well as the same number of non-crash cases from detector data, and used 45 minutes of detector data (in 1-min interval) preceding each crash in identifying the potential crash-leading traffic pattern. Six variables were considered: the mean and
variance (5-min) of occupancy, volume and speed. They found that, when considering only one variable, the distributions of each variable for both crash and non-crash cases 15 minutes prior to a crash are similar. That is, they were unable to identify the potential crash-leading pattern from the normal pattern based on one single variable. They also considered two variables and all the six variables with the combination of different time periods, however, none of the three methods could distinguish the crash-leading pattern from the normal non-crash pattern, and the overall classification error rate remained above 50%. They further explored the reason and suggested two possibilities: (1) limited sample size of the normal non-crash cases; (2) the spatial difference of the traffic flow, as they found that 68% of the crashes didn’t showing any significant change in the traffic flow patterns after the crash occurred.

In summary, research has investigated the prediction of incidents based on several parameters of operational conditions. In particular, one research investigated the prediction of one specific type of incidents - crash. The most often cited as significant factors in predicting incidents are occupancy, CVS, and standard deviation of speed. Other significant factors include standard deviation of volume, the speed difference between upstream and downstream, average speed, coefficient of variation of volume, standard deviations of time headway. Methods such as Bayesian analysis, log-linear models, binary logit models, and neural networks have been used for predicting incidents. Most of the literature do not divide traffic states, while two researches develops different models for non-congested and congested conditions. The literature review search did not reveal any applications of the PLM for predicting the
probability of incidents. There was also one research reporting that incidents cannot be predicted.

**Incident Verification**

Incident verification is one of the first steps in traffic incident management. It is the determination of the precise location and nature of an incident (Carson, 2009). It includes the verification of the time, location, and effect of an incident.

(1) Verification of Incident Time

Lee et al. (2003) developed a log-linear model using categorical variables for incident prediction, based on data of a 13-month period collected at 38 loop detector stations along the Gardiner Expressway in Toronto. They observed from data analysis that there was speed drop occurring when a queue formed after the crash occurrence and the backward-moving shock wave passed over the nearest upstream detector station. They then reported to consider the time when the speed abruptly dropped at the detector station immediately upstream of the crash site as the estimate of the actual time of crashes.

The determination of incident time was usually based on observations of occupancy vs. time series plots of the incidents, as shown by the research of Shaik (2003) and Masinick and Teng (2004). An example was shown in Figure 2-10.

Figure 2-10 showed whether a significant increase or decrease in occupancy was present in the travel direction. The beginning and ending points of the increase were visually determined for each incident, with the cutoffs representing the effective duration of the incident. However, the determination of the critical occupancy rate was not illustrated. There was no specific method to determine the value of occupancy increase
during an incident occurrence. Another difficulty was that it is hard to tell whether the increase in occupancy was caused by incident or recurrent congestion.

(2) Verification of Incident Location

This determination was based on the observation of sudden significant change in occupancy between upstream and downstream stations. Specifically, it was perceived that the station immediately upstream of each incident location would have the earliest and largest occupancy change, while occupancy change at subsequent upstream stations would start at a later time due to the backward moving shockwave (Masinick and Teng, 2004). An example was shown in Figure 2-11. The increase in occupancy at Station A was earlier than that at Station B and Station C, indicating that Station A was the immediate upstream station of the incident. However, multiple incident periods could give misleading results and thus need further treatment (Masinick and Teng, 2004). It is better to check the change in occupancy downstream, the flow and speed patterns as well.

(3) Verification of Incident Impact

After the verification of incident time and location, it is also necessary to determine whether an incident had significant impacts on capacity and traffic flow patterns. There was no previous literature regarding verifying the impact of incidents on freeway flow.

One research related to this verification was conducted by Masinick and Teng in 2004. He used a binary logit model to determine the rubbernecking likelihood of an incident and the significant variables that might cause rubbernecking. The results showed that four variables significantly influence whether an incident impact the traffic
in the opposite direction, they were peak hour, weather, presence of barriers, and weekend.

As reported by Carson (2009), the effectiveness of incident verification by loop detector data was very low to moderate. Thus several other methods of incident verification were also used in traffic incident management. For example, using on-site response personnel and remotely using closed-circuit television (CCTV), which have reported effectiveness of moderate to very high.

Summary

In conclusion, there was a significant amount of research on incident detection. Most of the algorithms are spatial measurement-based algorithms. Incidents are mainly detected based on occupancy. The California algorithms and McMaster algorithms are the most widely known and relatively simple compared to the later-developed algorithms, and are often used as a standard to evaluate the other algorithms. Most incident detection algorithms perform best in low-to-medium traffic volumes, while some work better in high volumes. The same algorithm might work differently at different geometric locations. One of the possibly serious problems in automatic incident detection is the "false alarm", in which an incident is signaled when one has not taken place. The literature review search did not reveal any applications of the PLM for the detection of incidents. The method or/and parameters often used in incident detection are summarized in Table 2-9.

Previous research on incident prediction is not so rich as on incident detection. Incidents are predicted from one or several parameters of operational conditions. There has also been research predicting the probability of one specific type of incident. The
most often used criteria in predicting incidents are occupancy and the variance of speed, as summarized in Table 2-10.

The time and location of incidents are mainly verified by occupancy. The time of incident occurrence can also be determined as the time when speed drops at the detector location immediately upstream of the incident location. Several other methods such as on-site response personnel and remotely using closed-circuit television (CCTV) are also used in incident verification. No previous research verifies the impact of incidents on operations to separate minor incidents from major incidents.

Future research should examine the following areas:

(1) Investigate whether incidents could be detected using other methods such as the PLM. It is recommended to divide traffic operation into different states, e.g., non-congested and congested conditions, and develop different algorithms for different conditions separately. Some other parameters such as the difference between speed limit and speed might be considered. Non-linear form of the parameters such as quadratic form might also be considered in detecting incidents.

(2) Insufficient data is also a weakness of previous research. Sufficient data from several freeways should be used to take into consideration the effect of geometric characteristics.

(3) Develop criteria of verifying the impact of incidents on operations. This method might help to avoid unnecessary analysis of minor incidents.

**Ramp Management Strategies Responsive to Incidents**

A large variety of incident management strategies had been developed to mitigate the impact of incidents on freeway operations. For instance, information dissemination with variable message signs (VMSs) or highway advisory radio, measures of reducing the clearance time of incidents, and operational strategies such as temporary opening of high-occupancy vehicle lanes to general traffic (Boyles et al., 2008). This section reviews literature related to ramp management strategies responsive to incidents, including ramp closure and ramp metering.
Ramp Closure

One commonly used strategy to improve traffic flow during incident conditions was to close on-ramps, which dates back to the early 1960s (Boyles et al., 2008). Miesse (1967) conducted a micro-simulation analysis to study the improvement attained from ramp closure. Michigan Department of Transportation conducted experiments in early 1960s on peak-hour ramp closure in Detroit, and found substantial increases in freeway throughput (up to 13.7%) and freeway speed (10 mph) without creating excessive problems on the arterials (Gervais and Roth, 1966).

Ritchie and Prosser (1990) developed a Freeway Real-Time Expert System Demonstration (FRED) for managing non-recurring congestion on urban freeways in Southern California. A real-time ramp metering, which can operate independently and only be interceded when the local capacity has been drastically reduced by an incident, was applied responsive to an incident in the FRED system. When an incident occurred, ramps upstream might be recommended for closure to reduce the demand at the incident site, and if the severity of the incident was above a specified threshold, all ramps within a certain distance upstream of the incident were recommended for closure.

Boyles et al. (2008) developed a two-phase procedure intended to guide the process of closing ramps in response to incidents. The first phase identified the best combination of on-ramps to be closed in response to an incident on a freeway section, using the total system travel time as measure of effectiveness. The second phase attempted to find the best length of time to close the ramps, using micro-simulation to study the vicinity of incidents in greater detail. The results showed that traffic operation was improved (traffic density is decreased) by closing the ramps during the incident.
However, their procedure was based on many assumptions, e.g., the incident duration was known.

**Ramp Metering**

Another ramp management strategy to respond to incidents is ramp metering. Chang et al. (1994) proposed an integrated real-time ramp metering algorithm, which could capture the dynamic traffic states with a two-segment linear flow-density model and thus provide time-varying metering rates, for non-recurrent freeway congestion. The effect of the incidents was considered by multiplying the mean flow rate with the incident factor that represents capacity reduction as a result of incidents. Metering rate was then obtained through optimization. The evaluation results showed that the proposed algorithm can increase freeway throughput, and the effectiveness increased with the severity of incidents and level of congestion.

Bogenberger et al. (2001) developed new adaptive fuzzy algorithms, named Adaptive and Coordinated Control of Entrance Ramps with Fuzzy Logic (ACCEZZ), and evaluated their effectiveness under incident scenarios for a particular 26 km freeway segment located on the northbound direction of Autobahn No. 9 (A9) in Munich, Germany by the microscopic simulator AIMSUN2. The severity (indicated by the number of lanes blocked) and the duration of the incidents were assumed to be detected immediately by an incident detection algorithm contained in the traffic management system and exactly estimated. The ACCEZZ models considered the effect of an incident through a suitable modification of the fundamental diagrams. These modifications included the implementation of special incident rates, incident capacities or desired incident occupancies, etc. Then the ramp metering rates of the upstream on-ramps were immediately adjusted. The evaluation results showed that the ramp...
metering substantially improved the freeway system performance by reducing the total
time spent in system.

The effects of existing ramp metering algorithms responsive to incidents were also
studied. Chu et al. (2004) evaluated the effectiveness of several ramp metering
algorithms under the non-recurrent incident scenario. They concluded that adaptive
ramp metering algorithms: ALINEA and BOTTLENECK, cannot improve system travel
time and freeway travel speed effectively under incident scenario, as the effectiveness
of ramp metering was marginal during severe congestion conditions. The coordinated
algorithm BOTTLENECK performed slightly better than the adaptive algorithms,
because the coordinated algorithm can respond to both local congestion and congestion
appealed in a coordinated area. They further stated that traveler-information systems
were important for reducing traffic congestion caused by incidents.

Michalopoulos et al. (2005) studied the effect of incidents on the performance of
Stratified Zone Metering (SZM) Strategy by setting artificial incidents in the microscopic
simulator AIMSUN. Each incident was assumed to cause blockage of lane(s) over a
certain period of time and occur on the rightmost lane at 17:00:00. The incident severity
was measured by duration of 10 minutes to 60 minutes with an increment interval of 10
minutes. The simulation results showed that the mainline, ramp and system total travel
time increased steadily as incident duration increases, and that an incident downstream
of a bottleneck was more detrimental to control performance than an upstream incident.
However, the authors didn’t provide adjustments of the SZM strategy in response to
incidents.
Ramp metering can also impact freeway operation safety. Lee et al. (2006) used a log-linear crash prediction model to investigate the effect of the local traffic-responsive ramp metering strategies on freeway safety. The results showed that ramp metering reduced crash potential by 5%-37% compared to the non-control case. The results provided some insight into how a local ramp metering strategy can be modified to improve safety (by reducing total crash potential) on longer stretch of freeways over a wide range of traffic conditions. However, the researchers only analyzed the effect of ramp metering control ALINEA on freeway safety, the effect of other algorithms on safety was not examined; Moreover, they used only one crash potential (speed difference) in the crash prediction model, and cannot reflect the relationship between crash and traffic flow accurately. Another limitation was that the microscopic simulation models used (PARAMICS) had not been sufficiently calibrated and the car-following behavior was not realistic.

Summary

Ramp closure is a ramp management strategy that is mainly used in response to the occurrence of incidents. Another approach is ramp metering. Generally, this method adjusts capacity under incident conditions by multiplying the mean flow rate with the incident factor that represents capacity reduction as a result of incidents (i.e., use a two-segment linear flow density model), or through a suitable modification of the fundamental diagrams. The corresponding metering rate is then adjusted or calculated from optimization. The evaluation results show that the adjustment can improve the system performance. Some existing ramp metering algorithms such as BOTTLENECK and ALINEA also show the ability or potential to improve system performance under
incident conditions. Moreover, ramp metering can also impact freeway operation safety by reducing crash potential.

Areas deserve further research include:

(1) To incorporate the capacity under incident condition and breakdown probability into the ramp metering algorithm. This application may delay or prevent the occurrence of breakdown and thus increase freeway throughput.

(2) Take into consideration the location of incidents in ramp metering. For instance, when incident occurs downstream of the bottleneck, the metering rate upstream should decrease, while when incident occurs upstream of the bottleneck, the metering rate might increase.

(3) Evaluate the effectiveness of ramp metering strategy that responsive to incidents using field data collected from several sites. This may help adjust and implement the proposed ramp metering strategy.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Definition of capacity</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persaud and Hurdle (1991)</td>
<td>Mean discharge flow</td>
<td>3 days data collected from a three-lane freeway</td>
</tr>
<tr>
<td>Agymang-Duah and Hual (1991)</td>
<td>2300 pc/h/ln under stable flow; 2200 pc/h/ln for post-breakdown conditions</td>
<td>Data of over 52 days during peak period</td>
</tr>
<tr>
<td>Minderhoud et al. (1997)</td>
<td>Non-congested flow data were used to estimate the capacity distribution</td>
<td>N/A</td>
</tr>
<tr>
<td>Cassidy and Bertini (1999)</td>
<td>Suggested the long-run queue discharge flow as the bottleneck capacity</td>
<td>3 days data collected from two freeway bottlenecks in Toronto</td>
</tr>
<tr>
<td>Lorenz and Elefteriadou (2001)</td>
<td>Flow rate corresponding to the expected probability of breakdown deemed acceptable</td>
<td>More than 40 congestion events occurring during about 20 days at two freeway bottlenecks in Toronto</td>
</tr>
<tr>
<td>Elefteriadou and Lertworawanich (2003)</td>
<td>Breakdown flow, maximum pre-breakdown flow, maximum discharge flow</td>
<td>More than 40 congestion events occurring during about 20 days at two freeway bottlenecks in Toronto</td>
</tr>
<tr>
<td>Kerner (2004)</td>
<td>Homogeneous segments: free / synchronized flow/wide moving jams Bottlenecks: determined by transition from free flow to synchronized flow, an infinite number of capacities Variable, considered only the volumes that cause breakdown as capacities</td>
<td>On highway A5-North in Germany</td>
</tr>
<tr>
<td>Brilon (2005)</td>
<td>The traffic volume below which traffic still flows and above which the flow breaks down, Weibull-distributed</td>
<td>Data of year 2000 collected from two freeway sections</td>
</tr>
<tr>
<td>Brilon et al. (2005)</td>
<td>Pre-queue flow and queue discharge flow</td>
<td>Data of year 2000 collected from two freeway sections</td>
</tr>
<tr>
<td>Banks (2006)</td>
<td>Maximum pre-breakdown flow, breakdown flow, maximum queue discharge flow and average queue discharge flow</td>
<td>18 extended data collection periods at 15 bottlenecks in North America</td>
</tr>
</tbody>
</table>
Table 2-2. Portion of freeway capacity available under incident conditions (Source: HCM 2010, HCM 2000)

<table>
<thead>
<tr>
<th>Number of freeway lanes by direction</th>
<th>Shoulder disablement</th>
<th>Shoulder accident</th>
<th>One lane blocked</th>
<th>Two lanes blocked</th>
<th>Three lanes blocked</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.95</td>
<td>0.81</td>
<td>0.35</td>
<td>0.00</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>0.99</td>
<td>0.83</td>
<td>0.49</td>
<td>0.17</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>0.99</td>
<td>0.85</td>
<td>0.58</td>
<td>0.25</td>
<td>0.13</td>
</tr>
<tr>
<td>5</td>
<td>0.99</td>
<td>0.87</td>
<td>0.65</td>
<td>0.40</td>
<td>0.20</td>
</tr>
<tr>
<td>6</td>
<td>0.99</td>
<td>0.89</td>
<td>0.71</td>
<td>0.50</td>
<td>0.26</td>
</tr>
<tr>
<td>7</td>
<td>0.99</td>
<td>0.91</td>
<td>0.75</td>
<td>0.57</td>
<td>0.36</td>
</tr>
<tr>
<td>8</td>
<td>0.99</td>
<td>0.93</td>
<td>0.78</td>
<td>0.63</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Table 2-3. Probability of lane closure due to crashes and breakdowns (Source: Chin et al., 2002)

<table>
<thead>
<tr>
<th>Type of crash</th>
<th>Number of vehicles involved</th>
<th>Lanes closed</th>
<th>No lane closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal crash</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 vehicle</td>
<td></td>
<td>0.892</td>
<td>0.108</td>
</tr>
<tr>
<td>More than 1 vehicle</td>
<td></td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Injury crash</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 vehicle</td>
<td></td>
<td>0.892</td>
<td>0.108</td>
</tr>
<tr>
<td>More than 1 vehicle</td>
<td></td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Property damage only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 3 cars and at most 1 truck</td>
<td>0.600</td>
<td>0.400</td>
<td></td>
</tr>
<tr>
<td>3 or more cars and/or 2 or more trucks</td>
<td>1.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Breakdowns</td>
<td></td>
<td>0.154</td>
<td>0.846</td>
</tr>
</tbody>
</table>

Table 2-4. Probability distribution of number of lanes closed (Source: Chin et al., 2002)

<table>
<thead>
<tr>
<th>Number of vehicles involved</th>
<th>Type of vehicles involved</th>
<th>1 Lanes closed</th>
<th>2 Lanes closed</th>
<th>3 Lanes closed</th>
<th>4+ Lanes closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Vehicle</td>
<td>Any type</td>
<td>0.997</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>2 Vehicles</td>
<td>2 cars, or 1 car and 1 truck</td>
<td>0.950</td>
<td>0.048</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>2 trucks</td>
<td>0.001</td>
<td>0.997</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>3 Vehicles</td>
<td>3 cars, or 2 cars and I truck</td>
<td>0.500</td>
<td>0.450</td>
<td>0.049</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>1 car and 2 trucks or 3 trucks</td>
<td>0.001</td>
<td>0.600</td>
<td>0.300</td>
<td>0.099</td>
</tr>
<tr>
<td>More than 3 Vehicles</td>
<td>Any type</td>
<td>0.001</td>
<td>0.099</td>
<td>0.800</td>
<td>0.100</td>
</tr>
</tbody>
</table>
Table 2-5. Capacity remaining after incidents (non-incident capacity = 1.000) (Source: Chin et al., 2004)

<table>
<thead>
<tr>
<th>Effect of crash</th>
<th>Number of freeway lanes</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle on the shoulder</td>
<td></td>
<td>0.450*</td>
<td>0.750</td>
<td>0.840</td>
<td>0.890</td>
<td>0.930*</td>
</tr>
<tr>
<td>1 lane blocked</td>
<td></td>
<td>0.000</td>
<td>0.320</td>
<td>0.530</td>
<td>0.560</td>
<td>0.750</td>
</tr>
<tr>
<td>2 lanes blocked</td>
<td></td>
<td>N/A</td>
<td>0.000</td>
<td>0.220</td>
<td>0.340</td>
<td>0.500</td>
</tr>
<tr>
<td>3 lanes blocked</td>
<td></td>
<td>N/A</td>
<td>N/A</td>
<td>0.000</td>
<td>0.150*</td>
<td>0.200*</td>
</tr>
<tr>
<td>4 lanes blocked</td>
<td></td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.000</td>
<td>0.100*</td>
</tr>
</tbody>
</table>

(* Assumed)

Table 2-6. Percent of freeway capacity available under incident conditions

<table>
<thead>
<tr>
<th>Time</th>
<th>Number of lanes</th>
<th>Lane closed shoulder</th>
<th>1 lane</th>
<th>2 lanes</th>
<th>3 lanes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goolsby, 1971</td>
<td>3</td>
<td>0.67</td>
<td>0.50</td>
<td>0.21</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.81</td>
<td>0.35</td>
<td>0.00</td>
<td>N/A</td>
</tr>
<tr>
<td>HCM 2010</td>
<td>3</td>
<td>0.83</td>
<td>0.49</td>
<td>0.17</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.85</td>
<td>0.58</td>
<td>0.25</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.87</td>
<td>0.65</td>
<td>0.40</td>
<td>0.20</td>
</tr>
<tr>
<td>Smith et al., 2003</td>
<td>3</td>
<td>N/A</td>
<td>0.37</td>
<td>0.23</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.75</td>
<td>0.32</td>
<td>0.00</td>
<td>N/A</td>
</tr>
<tr>
<td>Chin et al., 2004</td>
<td>3</td>
<td>0.84</td>
<td>0.53</td>
<td>0.22</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.89</td>
<td>0.56</td>
<td>0.34</td>
<td>0.15*</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.93*</td>
<td>0.75</td>
<td>0.50</td>
<td>0.20*</td>
</tr>
</tbody>
</table>

(* Assumed)
Table 2-7. Performance of some common incident detection algorithms (Sources: Subramaniam, 1991; Balke, 1993; Ozbay and Kachroo, 1999; Jeong et al., 2009)

<table>
<thead>
<tr>
<th>Type</th>
<th>Algorithm</th>
<th>Detection rate</th>
<th>False alarm rate</th>
<th>Mean time to detect (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparative</td>
<td>California Basic</td>
<td>82%</td>
<td>1.73%</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>California No. 7</td>
<td>67%</td>
<td>0.134%</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>California No. 8</td>
<td>68%</td>
<td>0.177%</td>
<td>3.04</td>
</tr>
<tr>
<td>Statistical</td>
<td>Standard Normal Deviate (SDN)</td>
<td>92%</td>
<td>1.3%</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Bayesian</td>
<td>100%</td>
<td>0%</td>
<td>3.9</td>
</tr>
<tr>
<td>Time series</td>
<td>Autoregressive Integrated Moving Average (ARIMA)</td>
<td>100%</td>
<td>1.5%</td>
<td>0.4</td>
</tr>
<tr>
<td>Smoothing</td>
<td>Autoregressive Integrated Moving Average (ARIMA)</td>
<td>100%</td>
<td>1.5%</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Double Exponential Smoothing</td>
<td>92%</td>
<td>1.87%</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>Filtering</td>
<td>95%</td>
<td>1.5%</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Low-pass Filter (LPF)</td>
<td>80%</td>
<td>0.3%</td>
<td>4.0</td>
</tr>
<tr>
<td>Traffic modeling</td>
<td>McMaster</td>
<td>68%</td>
<td>0.0018%</td>
<td>2.2</td>
</tr>
<tr>
<td>Artificial intelligent</td>
<td>McMaster</td>
<td>68%</td>
<td>0.0018%</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>Probabilistic Neural Networks (PNN)</td>
<td>89%</td>
<td>1.2%</td>
<td>0.9</td>
</tr>
<tr>
<td>Jeong et al. (2009)</td>
<td>Wavelet-based</td>
<td>95%</td>
<td>1%</td>
<td>1.34</td>
</tr>
</tbody>
</table>

Table 2-8. Estimated binary logit model for collision incident (Source: Balke et al., 2005)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Estimated Coefficient</th>
<th>t-ratio</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.508</td>
<td>-9.150</td>
<td>0.0000</td>
</tr>
<tr>
<td>Average Occupancy (%)</td>
<td>0.139</td>
<td>4.776</td>
<td>0.0000</td>
</tr>
<tr>
<td>Average CVS</td>
<td>14.251</td>
<td>3.703</td>
<td>0.0002</td>
</tr>
<tr>
<td>Number of observations</td>
<td>380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restricted log-likelihood</td>
<td>-212.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-186.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Algorithm name</td>
<td>Criteria used</td>
<td>Advantage / Disadvantage</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------</td>
<td>----------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>West (1969)</td>
<td>California algorithm 2</td>
<td>Occupancy from two adjacent detectors</td>
<td>Advantage: Simple, good performance and low false alarm rate; Disadvantage: cannot distinguish incident congestion and recurrent congestion</td>
</tr>
<tr>
<td>Payne and Tignor (1978)</td>
<td>California algorithm 7,8</td>
<td>Temporal differences in occupancy, downstream occupancy</td>
<td>Advantage: Simple, good performance and low false alarm rate; Disadvantage: cannot distinguish incident congestion and recurrent congestion</td>
</tr>
<tr>
<td>Persaud and Hall (1989)</td>
<td>McMaster algorithm</td>
<td>The pattern of flow, occupancy data</td>
<td>Advantage: effective at distinguishing traffic incidents from recurring traffic congestion; Disadvantage: requires a shift from uncongested to congested operation</td>
</tr>
<tr>
<td>Chassiakos and Stephanedes (1993)</td>
<td>Filter algorithm</td>
<td>Spatial occupancy difference between adjacent stations</td>
<td>Advantage: effective at distinguishing traffic incidents from recurring traffic congestion, excellent DR and FAR rates; Disadvantage: long detection time</td>
</tr>
<tr>
<td>Authors</td>
<td>Data / Tool used</td>
<td>Incident prediction criteria</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Liu (1997)</td>
<td>Incident data, traffic data, weather data</td>
<td>Over-heating vehicle incident: peak, merge, temperature, rain, speed variance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Binary Logit Model</td>
<td>Vehicle crash incident: merge, visibility, rain</td>
<td></td>
</tr>
<tr>
<td>Oh et al. (2001)</td>
<td>Accident data, real time traffic data</td>
<td>5-minute standard deviation of speed</td>
<td></td>
</tr>
<tr>
<td>Lee et al. (2003)</td>
<td>Categorical log-linear model</td>
<td>Lane by lane variation of speeds (CVS), traffic densities</td>
<td></td>
</tr>
<tr>
<td>Balke et al. (2005)</td>
<td>Integrated loop and weather data</td>
<td>Occupancy, average variation of speed</td>
<td></td>
</tr>
<tr>
<td>Pande et al. (2005)</td>
<td>Loop data. Logistic regression</td>
<td>5-min log CVS, standard deviation of flow, average occupancy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low-speed region: log of CVS and average occupancy, standard deviation of speed, average speed</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High-speed region: log of average occupancy, standard deviation of speed, average speed</td>
<td></td>
</tr>
<tr>
<td>Abdel-Aty et al. (2005)</td>
<td>Loop data. Logistic regression</td>
<td>CVS, occupancy, peak hour indicator, roadway indicator</td>
<td></td>
</tr>
<tr>
<td>Kuchangi (2006)</td>
<td>Loop data, incident log. Categorical log-linear model</td>
<td>Average occupancy, average variation of speed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rear-end crashes</td>
<td>Average speeds, average occupancy</td>
<td></td>
</tr>
<tr>
<td>Pande and Abdel-Aty (2006)</td>
<td>Loop data at the studied location and up to 1 mile upstream and downstream</td>
<td>Average speeds upstream and downstream of the station, difference in the lane occupancies across each individual lane</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2-1. Illustration of three parameters on time-series plot of flow and speed (Source: Elefteriadou and Lertworawanich, 2003)

Figure 2-2. Probability of F-S transition and the corresponding capacity ($q_{sum} = q_{in} + q_{on}$, that is, the mainline flow plus the on-ramp flow) (Source: Kerner, 2004)
Figure 2-3. Capacity distributions for a 3-lane freeway with and without variable speed control (13.5% average truck percentage, 5-minute interval) (Source: Brilon et al, 2005)

Figure 2-4. Probability of breakdowns in 15 min by Elefteriadou et al., reproduced (Source: Elefteriadou et al, 1995)
Figure 2-5. Probability of breakdown versus observed flow rate - Site “A” (Source: Lorenz and Elefteriadou, 2001)

Figure 2-6. Relationship between incident types and number of lanes affected. A) Incidents by type, B) number of lanes affected (N=17,885) (Source: Potter et al, 2007)
Figure 2-7. Relationship between v/c and accident rate at basic freeway section (Source: Chang et al., 2000)

Figure 2-8. Structure of basic California incident detection algorithm (Payne et al., 1976)
Figure 2-9. Predicted collision likelihood (Source: Balke et al., 2005)

Figure 2-10. Example of incident occupancy (Source: Masinick and Teng, 2004)
Figure 2-11. Incident-caused occupancy at multiple stations (Source: Masinick and Teng, 2004)
CHAPTER 3
DATA COLLECTION AND METHODOLOGY

As mentioned in chapter 1, the study on the impacts of incidents on freeway capacity and flow breakdown will depend largely on the analysis of data. This chapter illustrates the requirements of data types and collection sites, and presents the data analysis procedure.

Database Overview

Data collected from five freeway sections for the NCHRP 3-87 project “Proactive Ramp Management under the Threat of Freeway-Flow Breakdown” (Elefteriadou et al., 2009) are used for this dissertation. The data are described in Table 3-1.

Type of Data

At each site, three main types of data need to be obtained, which are traffic data, incident data, and weather data.

- Traffic data: include flow, occupancy and speed from all available detector stations along the selected sections. Traffic data are mainly collected from loop detectors and downloaded from corresponding websites.

- Incident data: include the information about the number of incidents, incident type and duration, etc., at the study site. They are collected from traffic data maintenance websites or incident logs developed by the local incident management center.

- Weather data: include average daily weather index such as the rain fall, snow, visual condition (fog, smoke, etc) at the data collection sites. The weather data are downloaded from the National Oceanic and Atmospheric Administration’s (NOAA) National Weather Service website (http://www.weather.gov/).

At some data collection sites, the geometric data and ramp management strategy (i.e., metering strategy and metering rate) can be collected. To cover the full change of traffic flow patterns and process of breakdown, the data are collected 24 hours daily or
at least from 6 AM to 7 PM. The data should also be available for about one year for the breakdown model development.

**Site Descriptions**

According to the objective of the NCHRP 3-87 project, the data collection site should be a freeway section that contains bottleneck, which is free from the effect of downstream bottlenecks and experiences recurrent congestion during the peak hour. This requirement is to make sure to obtain freeway capacity at non-incident conditions under free flow condition. A collection site may be several miles long, depending on the extent of congestion, and would likely encompass several merge and diverge areas. Moreover, wide geographic coverage within North America was attempted, to take into consideration the geographic effect on capacity and breakdown characteristics under incident conditions.

Based on these criteria, several sites are selected, which are Interstate 15 SB that is located in San Diego, California, Interstate 5 NB that is located in Sacramento, California, etc. The geographic characteristics, locations of loop detectors, and types of data available at each site are described below. The formats of data at each site are attached in Appendix A.

- Interstate 15 SB

  Interstate 15 SB is located in San Diego, California. The selected road section is from Balboa Avenue to I-52 with length of 2.42 miles. There are altogether 7 mainline detectors and 5 ramp detectors in this section. The configuration of the detector locations is illustrated in Figure 3-1. The most upstream detector is 1115802 and the most downstream detector is 1115779. The data collection period was from December, 2006 to November, 2007 (https://pems.eecs.berkeley.edu/). There are about 106 days
with incidents during the PM period through the data collection time at this site. The incident data includes information about the date, location, start time, duration, and causes of the incidents. Information on the number of lanes closed by incidents is not available.

- **Interstate 5 NB**

  Interstate 5 NB is located in Sacramento, California. The selected road section starts from Laguna Blvd to W. st with a length of about 10.41 miles. There are totally 16 mainline detectors along this section. The configuration of the detectors is illustrated in Figure 3-2. The data collection period was from November, 2006 to November, 2007. There are about 220 days with incidents during the AM period through the data collection time at this site. The incident data includes information about the date, location, start time, duration, and causes of the incidents. Information on the number of lanes closed by incidents is not available.

- **Queen Elizabeth Way (QEW, Toronto, Canada)**

  The selected road section is about 6.5 miles and consists of four interchanges, which are Cawthra Road interchange, Hurontario Street interchange, Mississauga Road interchange and Southdown road interchange. The most downstream interchange, the Cawthra Road interchange, was found to be free from downstream congestion and considered to be the bottleneck. The speed limit is 60 mi/h. A schematic of the section and the location of detectors are shown in Figure 3-3 and 3-4. The data are collected from the ICAT (ITS Centre and Testbed) platform website developed by Simon Foo at the University of Toronto. The data collection period was from January, 2005 to December, 2005. The metering control period is 6:00:00 AM – 10:00:00 AM. There are
about 294 incidents through the data collection time at this collection site. The incident data includes information about the date, location, start time, and causes of the incidents. Information on the number of lanes affected by incidents is available.

- **US 217 SB, Portland, Oregon**

  US 217 southbound is a 7-mile corridor that serves commuters during peak periods between downtown Portland and suburban areas in Beaverton, Tigard, Lake Oswego, etc. It diverges from US 26, intersects with Highways 8 (Canyon Rd.), 10 (Beaverton-Hillsdale Hwy), 210 (Scholls-Ferry Rd.), and 99W (Pacific Hwy), and finally merges onto I-5 southbound. A schematic of the section and the location of detectors are shown in Figure 3-5 and Figure 3-6. This freeway corridor contains 12 on-ramps, 10 of which are controlled by ramp meters. The ramp metering system on this freeway is supported by 36 loop detectors and 9 CCTV cameras, and uses the strategy SWARM since early November 2005. The data collection period was from Jan, 2006 to July, 2006 and Nov 2007 to Dec 2007. The incident data includes information about the date, location, start time, duration, type, and causes of the incidents. Information on the number of lanes affected is also available.

- **Interstate 494 EB, Minneapolis**

  Interstate 494 EB is located in Minneapolis. The selected road section is about 3 miles in length, from Bass Lake Rd to Rock Ford Rd. The locations of detectors are shown in Figure 3-7. The data collection period was from September, 2006 to August, 2007. There are about 930 incidents occurring through the data collection time at this site. The incident data includes information about the date, location, start time, and clear time of the incidents. Information on the number of lanes affected is also available.
Methodology

This subchapter presents data analysis procedure.

Data Screening

To guarantee the quality of data and accuracy of analysis, the collected data are screened in three aspects before further analysis, which are weather condition, incident condition and data quality.

Firstly, delete days with bad weather conditions to eliminate the impact of weather on capacity and traffic flow. Secondly, classify the data into several categories according to the time and location of incidents: condition with no incident, incidents occurring before congestion condition, incidents occurring during congested period condition, and incidents occurring downstream condition. Thirdly, data with poor loop detector performance are removed from the dataset.

(1) Weather Condition

Loop detector data are screened for bad weather days through the weather forecast obtained from local weather station. Days with bad weather conditions are excluded from the dataset. This is because adverse weather decreases the capacities and operating speeds on freeways, and thus will affect the accuracy of the data analysis results.

Previous research show that severe rain and snow cause the most significant reductions in capacities and operating speeds. The HCM 2010 addresses the impact of adverse weather on freeway capacity. The reported reductions in capacity caused by heavy snow are 19.5-27.8%. The reported reductions in capacity caused by heavy rain are 10.7%-17.7%. Agarwal et al. (2005) concluded from four years data analysis that heavy rains (more than 0.25 inch/hour) and heavy snow (more than 0.5 inch/hour)
reduce capacity by respectively 10%-17% and 19%-27%, and reduce speed by respectively 4%-7% and 11%-15%. Generally, these values are significantly lower than those specified by the HCM 2000. They also found that even trace, light snow cause 3%-5%, 7%-9% reduction in speed and 3%-5%, 6%-11% reduction in capacity.

In a word, adverse weather condition leads to different traffic flow performance compared to the clear weather. Thus, nearly all traffic engineering guidance and methods that are used to estimate highway capacity assume clear weather conditions (Agarwal et al., 2005).

In this dissertation, adverse weather is viewed as:

- Rain fall > 0.20 inch/day, or/and
- There is snow dropping during the day, or/and
- There is heavy fog, hail, blowing snow and tornado conditions.

(2) Incident Condition

After the screening of adverse weather data, traffic data at the same site are separated into two categories: without incidents and with incidents occurring. Data with incidents and some data with no incident (selected to be close to the date of incidents) are kept for further analysis.

Moreover, incidents are grouped into three categories by time and location: incidents occurring before congestion, incidents occurring during the congested period, and incidents occurring downstream of the study site. Each category could be further classified by occurring at the bottleneck and non-bottleneck locations.

(3) Data Quality

Good quality data is important for freeway traffic control algorithms and analysis. There are many data screening criteria in previous research. For example, Cleghorn et
al. (1991) screened the data, which were collected from the Queen Elizabeth Way (QEW) in Ontario, Canada, based on the following criteria: a) volume must not be negative and must not exceed 60 vehicles per interval; b) Occupancy must not be negative and must not exceed 100 percent; c) Speed must not be negative and cannot exceed 150 km/h (93 mph). The screening test is performed every interval (30 seconds) and a flag is set to mark the invalid datum once any of the tests failed.

For data collected for this dissertation, a flag value exists in the data, which identifies the quality of the data. When the value of the flag is 0, then the data collected is good in quality. Otherwise, the data is not good. Therefore, data with the flag value non-zero are viewed to be bad and excluded from the dataset. Moreover, some websites have pointed out which detectors may be bad functioning during which time period. This may also help remove the bad quality data.

**Incident Verification**

In most of the cases, the reported incident time was a few minutes later than the actual incident time. Moreover, sometimes an incident may have no or little influence on traffic flow due to the demand, time and incident characteristics. Thus it is necessary to verify the actual time, location and impact of incidents from various sources. These verifications are important to the accuracy of incident analysis, and can help avoid unnecessary analysis of minor incidents.

Based on the literature review, this dissertation verified the incidents from occupancy, speed, and the time-series plots of speed and flow. If there was obvious occupancy and speed change caused by the incident, and if there were large change in the time-series plots pattern, the incident was then considered to be significant in effect and kept for further analysis. Otherwise, the data would be removed.
As it is difficult to distinguish incident-induced congestion from recurrent congestion, the differences between occupancy changes caused by incidents and that caused by congestion should be applied in verifying incidents. Firstly, apart from occupancy, it also need to check the flow and speed during the incidents. Usually, incidents may cause a short-period occupancy increase relative to recurrent congestion. Secondly, check the occupancy upstream and downstream of the incident location. When an incident occurs, occupancy will increase at the upstream location and decrease at the downstream location. While when congestion experienced, occupancy will propagate from the downstream bottleneck to upstream, and there is usually no much decrease in occupancy at the downstream locations.

Based on the data analysis, an incident that is verified to affect traffic has the following characteristics:

- Occupancy immediately upstream of the incident location is greater than 20 percent: \( \text{occ}_i > 20 \);  
- Occupancy immediately downstream of the incident location is less than 10 percent: \( \text{occ}_{i+1} < 10 \) (except for incidents occurring at or downstream of the most downstream detector location, as there is no downstream data collected);  
- The change in occupancy and speed lasts more than 10 minutes.

An example from data of January 28, 2005 at the preliminary data analysis site is illustrated to show the verification process of incidents. On January 28 2005, the incident log indicates that there was a collision occurring at 15:05 at Cawthra Rd (downstream of detector 480DES). Time series plot of occupancy is made around 15:00 at the detector 480DES as well as the upstream detector 470DES and the downstream detector 500DES, as shown in Figure 3-8 A). It shows that when the incident occurs, occupancy at the upstream detector (470DES and 480DES) generally increase from
10% to 50%, while occupancy at the downstream detector (500DES) decreases from 10% to about 4%. Moreover, occupancy at the detector just upstream of the incident location, i.e., 480DES, increases earlier and also ends earlier than 470DES. Figure 3-8 B) describes the time series flow and speed plots at detector 480DES. It shows that, there was congestion at 15:00 - 15:55 caused by the incident. The speed and flow dropped sharply to a very low level during the incident and recovered to a higher level after the incident. Thus, the incident was verified to occur at 15:00-15:30 downstream of the detector 480DES, and that it has effect on traffic flow.

Data Analysis Procedure

Data are analyzed through the following procedure,

(1) All the data are firstly screened for weather and data quality. Secondly, the time, location and effects of incidents are verified from occupancy, speed, and time series plots. For incidents verified to disrupt traffic, the speed, occupancy, and flow as well as their variances at the beginning of incidents are compared. The impact of incidents on operational conditions is analyzed from time-series plots and density maps.

(2) Define breakdown occurrence criteria. In this dissertation, a breakdown is defined as five or more consecutive 1-minute intervals with speeds drops by 10 mph. This was established based on the time series speed plots at each detection location for several days. The breakdown is viewed to be activated if the following three criteria are met (Elefteriadou et al., 2009).

i. Speed differences between two consecutive minutes is negative:
   \[ \Delta S_i = S_i - S_{i-1} < 0 \]

ii. The average speed during the previous 5 minutes is greater than the average speed in the following 5 minutes by at least 10 mph (or 16 km/h):
   \[ \text{Avg}\{S_{i-5},...,S_{i-1}\} > \text{Avg}\{S_{i},...,S_{i+4}\}+10 \text{mph} \]

iii. The maximum speed during the following 10 minutes (default value) is less than the speed before the speed drop.
   \[ \text{Max}\{S_{i},...,S_{i+9}\} < S_{i-1} \]

To avoid confusion, this dissertation names recurrent congestion as demand-induced breakdown, and the congestion caused by incidents as incident-induced breakdown.
(3) Analyze the relationship between incidents and beginning of congestion. Further investigate the potential of developing a probabilistic incident-induced breakdown model.

(4) A definition and measurement of freeway capacity for incident conditions is proposed. Six parameters are proposed to define capacity for incident condition, which are respectively defined as,

- **Maximum pre-breakdown flow** is the highest flow that occurred within 10 minutes before the breakdown or incident-induced breakdown during non-congested conditions.

- **Breakdown flow** is the one-minute flow per lane for the interval before the demand-induced breakdown or incident-induced breakdown (i.e., before the speed drops by 10 mph for 10 minutes).

- **Average flow for 10 minutes before breakdown** is the average of ten minutes flow per lane before the breakdown.

- **Average discharge flow during entire congestion** is the average flow per lane during the congested conditions.

- **Average discharge flow during both incident and congestion** is the average flow per lane during both the incident conditions and congested conditions.

- **Minimum 10-minute flow rate** is the minimum 10-minute flow rate by lane (by moving average of ten successive 1-minute flows to reduce variation) measured in the bottleneck created by an incident.

(5) Investigates the relationship between incident probability and operational conditions. Observations from the preliminary data analysis will be used in this step.

**Database Overview**

This subsection presents an overview of the database. The number of days of data analyzed for the dissertation is summarized in Table 3-2 by different traffic conditions and locations.
### Table 3-1. Description of data available

<table>
<thead>
<tr>
<th>Site</th>
<th>State</th>
<th>Length (mi)</th>
<th>Traffic data</th>
<th>Weather data</th>
<th>Incident data</th>
<th># Lanes affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-15 SB</td>
<td>San Diego, CA</td>
<td>2.4</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>I-5 NB</td>
<td>Sacramento, CA</td>
<td>10.4</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>QEW</td>
<td>Toronto, Canada</td>
<td>6.5</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>US 217</td>
<td>Portland, OR</td>
<td>7</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>I-494 EB</td>
<td>Minneapolis, MN</td>
<td>3</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

### Table 3-2. Summary of data analysis

<table>
<thead>
<tr>
<th># Data points (days)</th>
<th>Data with no incident at 'B'</th>
<th>Incident before congestion at 'B'</th>
<th>Incident during congestion at 'B'</th>
<th>Incident downstream at 'B'</th>
<th>Incident at 'O'</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto</td>
<td>56</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>34</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>59</td>
</tr>
<tr>
<td>Portland</td>
<td>30</td>
<td>3</td>
<td>12</td>
<td>8</td>
<td>8</td>
<td>67</td>
</tr>
<tr>
<td>San Diego</td>
<td>73</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>87</td>
</tr>
<tr>
<td>Sacramento</td>
<td>65</td>
<td>10</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>99</td>
</tr>
<tr>
<td>Sum</td>
<td>258</td>
<td>16</td>
<td>29</td>
<td>36</td>
<td>16</td>
<td>8</td>
</tr>
</tbody>
</table>

*'B' indicates bottleneck location, 'O' indicates other locations.*
Figure 3-1. Location of detectors at I-15 SB

Figure 3-2. Location of detectors at I-5 NB

Figure 3-3. Schematic of detectors at the Queen Elizabeth Way (QEW) Site
Figure 3-4. Location of detectors at QEW

Figure 3-5. Schematic of the Portland site

Figure 3-6. Location of detectors at the Portland site
Figure 3-7. Location of detectors at the Minneapolis site

Figure 3-8. Verification of incident on Jan 28, 2005. A) Change in occupancy, and B) Change in flow.
CHAPTER 4
QUALITATIVE ANALYSIS OF THE IMPACTS OF INCIDENTS ON OPERATIONAL CONDITIONS

Methods often used in freeway bottleneck analysis include time-series plots, density maps, and cumulative curves of vehicle counts (N-curves). This chapter analyzes the impacts of incidents on operational conditions qualitatively based on the first two types of analysis. The third method is not considered, as the N-curve method is vulnerable to the accumulation of errors resulting from biased counts and its contribution is modest when compared with that of careful speed analysis, as reported by Kurada et al. (2007). For freeway segments with many on-ramps and off-ramps, errors in vehicle counting are difficult to avoid due to reasons such as unstable detector performances. Thus, the N-curves at complex freeway bottlenecks would not be the best approach. This qualitative analysis can provide a general idea and insights on achieving the stated objectives.

Time-series Plots

This subchapter analyzes the relationship between incidents and the beginning of congestion.

(1) Conditions with no incident

Traffic data from January 12, 2005 at the Toronto site is used as an example of condition with no incident, as shown in Figure 4-1. Figure 4-1 A) describes the change of occupancy around congestion: occupancy increases earliest at the most downstream detector 510DES and then at the upstream detectors. Figure 4-1 B) describes the change of speed at the beginning of congestion. A speed drop of 12 mi/h is observed firstly at the detector 510DES at 6:20 AM. These observations indicate that the location of the most downstream detector 510DES is the bottleneck. When comparing Figure A)
and B), change in speed is relatively larger than the change in occupancy at the beginning of congestion. Figure 4-1 C) shows the flow and speed before, during, and after congestions from 6:22-11:05 at the breakdown location.

Similar patterns in the time series plots exist during other recurrent congested periods. Breakdown starts at the bottleneck usually during AM peak periods. When breakdown occurs, speed drops first at the bottleneck location. Congestion propagates to upstream detectors, thus, congestion at the upstream detectors starts later and ends earlier.

(2) Incidents occurring before congestion

Traffic data from March 9, 2005 at Toronto site is used as an example of incidents occurring before congestion conditions. The incident log indicates that there is an incident occurring just downstream of the detector 480DES at 17:32. Thus, data at the detector 480DES, the upstream detector 470DES and the downstream detector 500DES are analyzed. There is no breakdown during the PM peak period this day.

Figure 4-2 A) shows that the occupancy upstream of the incident location (470DES, 480DES) increases from about 10% to 50%, while occupancy downstream of the incident location (500DES) decreases from about 10% to 5%. This is different from the change of occupancy at the beginning of congestion for condition with no incident, during which the downstream occupancy does not change much. Figure 4-2 B) shows that the speed upstream of the incident location (470DES, 480DES) decreases from about 65 mi/h to 10 mi/h, while speed downstream of the incident location (500DES) increases from about 65 mi/h to 75 mi/h. Occupancy and speed change earliest at the
detector 480DES, e.g., there is a speed drop of 20 mi/h at 17:17 at 480DES. Thus, the incident was observed to be at 480DES during 17:17-18:01.

Figure 4-2 C) shows the speed and flow at the detector 480DES before, during, and after the incident. It is observed that there is congestion caused by the incident at 480DES during 17:18-18:49. There are two operational conditions caused by the incident during congestion: during both incident and congestion (17:17-18:01), after incident and during congestion (18:01-18:49). The average flows during the two operational conditions are respectively 899 veh/h/ln and 1835 veh/h/ln. The average speeds during the two operational conditions are 9.4 mi/h and 36.9 mi/h.

In summary, changes in flow and speed caused by incidents are much steeper than that caused by congestion, and last a relatively short period of time.

(3) Incidents occurring during the congested period

Traffic data from Oct 6, 2005 at Toronto site is used as an example of incidents occurring during congested conditions, as shown in Figure 4-3. The incident log indicates that there is an incident occurring just downstream of the detector 440DES at 8:12. Thus data at the detector 440DES, the upstream detector 420DES and the downstream detector 460DES are analyzed.

Figure 4-3 A) shows that the occupancy upstream of the incident location (420DES, 440DES) increases from about 30% to 60%, while occupancy downstream of the incident location (460DES) decreases from about 30% to 5%. Detector 440DES is the first to show an increase in occupancy at 7:58. Figure 4-3 B) shows that the speed upstream of the incident location (420DES, 440DES) decreases from about 25 mi/h (already in congestion) to 8 mi/h, while speed downstream of the incident location
Detector 440DES first shows a speed drop of 14.5 mi/h at 7:58. Based on the observation from Figure 4-3 A)-B), the incident was verified to be at 8:02-8:38 at detector 440 DES.

There is breakdown at the bottleneck 510DES at 6:33. Congestion propagates to the upstream detector 440DES at 6:37, and recovers at 9:43. Figure 4-3 C) shows time series plots of speed and flow immediately upstream of incident location at 440DES. Speed and volume drop sharply after the incident. The average discharge flow was 485 veh/h/ln during the incident, while 1709 veh/h/ln after the incident (during the congested period). The average speed was 9.2 mi/h during the incident, while 25.2 mi/h after the incident. From this viewpoint, the incident aggravated the congestion.

(4) Incidents occurring downstream of the bottleneck

Traffic data from Sep 19, 2005 at Toronto site is used as an example of incidents occurring downstream of the bottleneck. The incident log indicates that an incident occurs at 12:29 at Dixie Rd (downstream of 510DES). Thus data at detector 510DES, and at the upstream detectors 500DES and 480DES, are analyzed. There is no downstream detector data available.

Figure 4-4 A) shows that occupancy increases greatly after the incident at all the detectors (all upstream of the incident). Figure 4-4 B) shows that detector 510DES has a speed drop of 55 mi/h first after the incident at 12:27. Figure 4-4 C) shows the time-series plots of flow and speed at the detector 510DES, which indicates that there is congestion caused by the incident from 12:27-13:25.

The beginning of congestion for incidents occurring downstream conditions has similar pattern to that for incidents occurring before congestion condition, except that
the impact of incidents on operations might be mitigated by the distance of the incidents from the bottleneck.

(5) Minor Incidents

Some incidents are identified to have little effect on traffic flow. They may cause the upstream occupancy increasing for a short time period, but not effective enough to cause or aggravate congestion. In this case, flow and speed immediately upstream of the incident location may drop sharply for a short period of time, however, the impact on flow at further upstream detectors is little. Data of November 8, 2005 at Toronto site is used as an example.

The incident log indicates that an incident occurs at 21:18 near the detector 510DES. The incident was observed to be at 510DES at 21:19-21:31 based on the change of occupancy and speed, as shown in Figure 4-5 A) and B). Figure C) shows that, when the incident occurs, the average speed at 510DES drops to 17 mi/h and the average volume drops to 723 veh/h/ln However, the disruption in traffic flow only lasts 13 minutes.

Conclusions Related to the Time Series Plots

The number of days examined, the start time and duration of congestion under non-incident and various incident conditions are summarized in Table 4-1. The third column provides the number of days for each condition. The fourth column provides the range of incident duration. The fifth column indicates that the start time of congestion is different for non-incident and incident conditions: congestion might occur at any time through the day for incident conditions while is concentrated in the peak periods for non-incident conditions. The sixth column indicates that the congestion durations for
incidents occurring before congestion and downstream of the bottleneck are relatively shorter than the other two conditions.

A comparison of the changes in operational conditions for normal (5-minute before and after breakdown) and incident conditions (5-minute before and after incident) are summarized in Table 4-2. The distributions of the changes for non-incident and incident conditions are illustrated in Figure 4-6.

Operations can be classified into four states by the time and duration of incidents: before congestion (10-minute), during incidents and congestion, after incidents and during congestion, post congestion recovery (10-minute). The average values of flow, speed, and occupancy during different operational states for various incident conditions are respectively shown in Column 3 to Column 6 in Table 4-3. San Diego and Sacramento sites have few crashes verified and thus are not included in Table 4-3.

Six conclusions on the impact of incidents on operational conditions are drawn:

(1) Demand-induced breakdown starts at the bottleneck usually during the AM peak periods. When breakdown happens, congestion propagates to upstream detectors, thus congestion at the upstream detectors starts later and ends earlier.

(2) Incident-induced breakdown might occur at both bottleneck and non-bottleneck locations, during the peak and off-peak periods. Detectors immediately upstream of the incident location are the first to show an abrupt increase in occupancy, while detectors immediately downstream show a decrease in occupancy, similar to the observations made by Shaik (2003) and Masinick and Teng (2004).

(3) Apart from occupancy, speed can also be used in verifying incidents. Detectors immediately upstream of the incident location are the first to show an abrupt decrease in speed, while detectors immediately downstream show an increase in speed. Changes in speed at the beginning of congestion are relatively more obvious than the changes in occupancy.

(4) Changes in operational condition at the beginning of congestion are much steeper for incident conditions than for non-incident conditions. For instance, at the Toronto site, the average 5-minute changes in flow, speed, and occupancy before/after incidents are respectively -31%, -65%, and 218%; while the average
5-minute changes in flow, speed, and occupancy before/after congestion (for non-incident conditions) are respectively -3%, -32%, and -32%.

(5) Comparing the 10-minute operations before congestion with the 10-minute operations after recovery (as shown in Table 4-3), there is not much difference between the average speed and occupancy. However, average flow 10-minute before congestion is generally much higher than the average flow 10-minute after recovery.

(6) Comparing congested conditions during/after incidents (as shown in Table 4-3), incidents aggravated the congestion. Take the Toronto site as an instance, the average flow, occupancy, and speed during incidents and congestion are around 950 veh/h/ln, 40, and 16 mi/h; The average flow, occupancy, and speed after incidents and during congestion are around 1700 veh/h/ln, 25, and 35 mi/h.

Density Maps

Shock wave analysis is usually used in travel time estimation and travel demand analysis, and might be used in estimating incident duration. This subchapter compares the density maps under non-incident and different types of incident conditions, to illustrate the impact of incidents on congestion propagation. Density is calculated based on the occupancy, detector length and vehicle length using equation 4-1:

\[
k = \frac{52.8}{L_v + L_D} \left( \frac{\% OCC}{18 + 6} \right)
\]

(Eq. 4-1)

Where, the vehicle length \( L_v \) and detector length \( L_D \) are respectively 18 feet and 6 feet. The calculation time step is every 5 minute or 1 minute, according to the duration of incidents and congestion. Flow with density less than 40 vehicles per lane-mile is viewed to be un-congested flow, flow with density between 40 and 65 is viewed to be unstable flow that is near-capacity flow condition, while flow with density greater than 65 is viewed to be congested flow (May, 1990).

Density maps are only plotted for the Toronto site, as Minneapolis and Portland sites have two or more than two bottlenecks with one in the middle of the road section,
which make it difficult to identify the reason of congestion. In the density maps below, the horizontal axis are the detectors along the travel direction with 510DES or 480DES as the bottleneck, the vertical axis is time. A five star indicates the location and start time of an incident. Shock waves are viewed to be the density curves of 65 vehicles per lane-mile.

(1) Condition with no incident

Data of February 10, 2005 is used as an example of condition with no incident. There was breakdown at 6:31-10:59 at the bottleneck (detector 510DES). Figure 4-7 shows that the shockwave forms a regular pattern with mainly three shockwaves: a frontal stationary shockwave at the bottleneck, a backward forming shock wave, and a forward recovery shockwave. The highest densities were only little above 100 and occur at only few detector locations.

The congestion propagates to the most upstream detector 370DES at 7:25, taking 49 minutes. The distance between 370DES and 510DES is about 10.25 km. Thus, the speed of the shockwave is 7.8 mi/h.

(2) Incidents occurring before congestion

Data of March 9, 2005 and May 25, 2005 are used as examples of incidents occurring before congestion conditions, as shown in Figure 4-8 and 4-9. On March 9 2005, an incident occurred downstream of the detector 490DES at 17:17-18:01, and caused incident-induced breakdown at 17:18-18:03. On May 25 2005, an incident occurred downstream of the detector 440DES at 9:57-10:26, and caused incident-induced breakdown at 9:59-10:27. It is observed from Figures 4-8-4-9 show that the frontal stationary shock wave occurred immediately upstream of the incident locations.
but not at the bottleneck, and that the shock wave recovers shortly following the clearance of the incident. Contrary to the non-incident conditions, density downstream of the incident location is very low during the study period, and the highest density (greater than 100) appeared at more detector locations.

The speeds of shockwaves are calculated. On March 9, congestion propagates to the most upstream detector 370DES at 17:59, taking 42 minutes. The distance between the incident location 480DES and 370DES is about 7.8 km. Thus, the speed of the shockwave is 6.9 mi/h. On May 25, congestion propagates to the most upstream detector 370DES at 10:16, taking 19 minutes. The distance between the incident location 440DES and 370DES is about 4.8 km. Thus, the speed of the shockwave is 9.5 mi/h.

(3) Incidents occurring during the congested period

Data of October 6, 2005 is used as an example of incidents occurring during congested period conditions. There is congestion at 6:37-9:43 at the most downstream detector 510DES. An incident occurred at 8:02-8:38 at detector 440DES. Figure 4-10 shows that the shockwaves have similar shape as that for non-incident conditions (Figure 4-7). However, during the incident, upstream density increased greatly while downstream density decreased sharply. The highest densities are high above 100 while the lowest densities approaching 5.

Figure 4-11 illustrates the change in density during the incident in detail with 1-minute time step. The shockwave with density greater than 100 propagates to the most upstream detector 370DES at 8:11, taking 9 minutes. The distance between the incident
location 440DES and the location at 370DES is 4.83 km. The speed of the shockwave is 16.4 mi/h.

(4) Incidents occurring downstream

Data of September 22, 2005 is used as an example of incidents occurring downstream, as shown in Figure 4-12. There was an incident occurring at 17:16-19:46 downstream of the most downstream detector 510DES. Figure 4-12 shows that the shockwaves cover relatively a small area. The two downstream detectors 510DES and 500DES have almost the highest density with the longest duration.

Congestion propagates to the upstream detector 380DES at 17:40, taking 25 minutes. The distance between the incident location 510DES and the location at 380DES is 9.33 km. The speed of the shockwave is 13.9 mi/h.

Conclusions Related to Density Maps

When congestion (non-incident conditions) or an incident begins, densities might change with time and are different at the upstream and downstream locations. Percentage changes in densities at the beginning of congestion or incidents at Toronto site are summarized in Table 4-4, based on analysis of the density maps illustrated above.

Comparing the density maps for different conditions, several preliminary conclusions are drawn regarding the impact of incidents on congestion,

(1) For non-incident conditions (e.g., see Figure 4-6), the density map forms a regular shape with mainly three shockwaves: a frontal stationary shockwave, a backward forming shock wave, and a forward recovery shockwave. The frontal stationary shockwave occur at the bottleneck. The highest densities were only little above 100 and occur at only few detector locations.

(2) For incidents occurring before congestion (e.g., see Figures 4-7 and 4-8), the frontal shockwave moves from the bottleneck to the incident location. The shock
wave usually recovers following the clearance of the incident. The highest density (greater than 100) appeared at more detector locations.

(3) For incidents occurring during congested period (e.g., see Figure 4-9), the shockwaves have similar shapes to that at non-incident conditions. However, during the incident, density upstream of the incident location increased greatly while density downstream decreased sharply. The highest densities are high above 100 forming a rectangle, with the lowest densities approaching 5.

(4) For incidents occurring downstream (e.g., see Figure 4-10), the shockwaves concentrate at the most downstream detectors and have a relatively smaller range, probably due to the distance between the incident location and the bottleneck.

(5) The 5-minute changes in density caused by incidents (-36.45%) are larger than that caused by recurrent congestion (-10.32%) (as shown in Column 3 of Table 4-4). The downstream densities decrease by 76% when incidents occur, while do not have change under non-incident conditions. However, these observations are based on a limited amount of data points.

(6) Based on the analysis of density maps illustrated above, the speeds of shockwave are 7.8 mi/h and 6.4 mi/h for non-incident conditions, 6.9 mi/h and 9.5 mi/h for incident occurring before congestion conditions, 16.4 mi/h for incidents occurring during congested period conditions, and 13.9 mi/h for incident occurring downstream conditions. Generally, the speeds of shockwave for incident conditions are larger than that for non-incident conditions. However, these observations are based on a limited amount of data points.

Conclusions

A qualitative analysis of the impact of incidents on operational conditions is conducted through time series plots and density map. Speed seems to be more appropriate than occupancy in incident verification. There is reduction in discharge flow caused by incidents. Generally, the speed of shockwaves for incident conditions is larger than that for non-incident conditions. It is suggested that comparing the changes in operational conditions at the beginning of congestion for non-incident and incident conditions can help predict the probability of incidents and incident-induced breakdown.
Table 4-1. Relationship between incidents and beginning of congestion

<table>
<thead>
<tr>
<th>Site</th>
<th>Type of data</th>
<th># Days</th>
<th>Incident duration range (min)</th>
<th>Start time of congestion</th>
<th>Congestion duration range (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto</td>
<td>No incident</td>
<td>56</td>
<td>N/A</td>
<td>5:47- 7:31</td>
<td>62 - 360 (182)*</td>
</tr>
<tr>
<td></td>
<td>Before congestion</td>
<td>11</td>
<td>16 -80 (35)</td>
<td>6:30- 19:30</td>
<td>19 - 266 (71)* *</td>
</tr>
<tr>
<td></td>
<td>During congestion</td>
<td>9</td>
<td>25 - 41 (32)</td>
<td>6:30- 16:00</td>
<td>56- 237 (153) *</td>
</tr>
<tr>
<td></td>
<td>Downstream</td>
<td>4</td>
<td>25 - 59 (36)</td>
<td>9:45- 22:30</td>
<td>26 - 58 (36) *</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>No incident</td>
<td>35</td>
<td>N/A</td>
<td>6:50-7:51</td>
<td>9-138 (67)</td>
</tr>
<tr>
<td></td>
<td>Before congestion</td>
<td>16</td>
<td>12-83 (42)</td>
<td>5:40-18:30</td>
<td>22-107 (51)</td>
</tr>
<tr>
<td></td>
<td>During congestion</td>
<td>5</td>
<td>31-62 (44)</td>
<td>6:50-8:05</td>
<td>35-121 (72)</td>
</tr>
<tr>
<td></td>
<td>Downstream</td>
<td>1</td>
<td>77</td>
<td>8:13</td>
<td>78</td>
</tr>
<tr>
<td>Portland</td>
<td>No incident</td>
<td>33</td>
<td>N/A</td>
<td>14:38-16:29</td>
<td>54-385 (173)</td>
</tr>
<tr>
<td></td>
<td>Before congestion</td>
<td>20</td>
<td>20-91 (47)</td>
<td>3:44-19:00</td>
<td>24-117 (63)</td>
</tr>
<tr>
<td></td>
<td>During congestion</td>
<td>13</td>
<td>20-104 (42)</td>
<td>7:20-19:00</td>
<td>37-328 (136)</td>
</tr>
<tr>
<td></td>
<td>Downstream</td>
<td>1</td>
<td>58</td>
<td>8:06</td>
<td>75</td>
</tr>
</tbody>
</table>

*Note: * Indicates the average value.

Table 4-2. Changes in operational conditions 5-min before/after breakdown or 5-min before/after incidents

<table>
<thead>
<tr>
<th>Percent changes</th>
<th>Flow No incident</th>
<th>Flow Incident</th>
<th>Speed No incident</th>
<th>Speed Incident</th>
<th>Occupancy No incident</th>
<th>Occupancy Incident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto</td>
<td>Min. 28.00%</td>
<td>-90.00%</td>
<td>17.00%</td>
<td>19.00%</td>
<td>-92.00%</td>
<td>1761.00%</td>
</tr>
<tr>
<td></td>
<td>Avg. 3.48%</td>
<td>31.68%</td>
<td>32.44%</td>
<td>64.85%</td>
<td>-46.13%</td>
<td>-217.71%</td>
</tr>
<tr>
<td></td>
<td>Max. 21.00%</td>
<td>65.00%</td>
<td>53.00%</td>
<td>89.00%</td>
<td>-21.00%</td>
<td>-32.00%</td>
</tr>
<tr>
<td></td>
<td>Min. -13.00%</td>
<td>-34.83%</td>
<td>18.82%</td>
<td>11.34%</td>
<td>-174.75%</td>
<td>-539.47%</td>
</tr>
<tr>
<td></td>
<td>Avg. 8.76%</td>
<td>32.60%</td>
<td>35.20%</td>
<td>54.43%</td>
<td>-54.07%</td>
<td>-158.85%</td>
</tr>
<tr>
<td></td>
<td>Max. 51.68%</td>
<td>100.00%</td>
<td>65.76%</td>
<td>100.00%</td>
<td>17.51%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Portland</td>
<td>Min. -25.40%</td>
<td>-15.99%</td>
<td>24.32%</td>
<td>-12.90%</td>
<td>-131.60%</td>
<td>-295.73%</td>
</tr>
<tr>
<td></td>
<td>Avg. 11.32%</td>
<td>21.69%</td>
<td>46.17%</td>
<td>46.00%</td>
<td>-73.43%</td>
<td>-98.10%</td>
</tr>
<tr>
<td></td>
<td>Max. 36.45%</td>
<td>64.43%</td>
<td>66.12%</td>
<td>87.26%</td>
<td>57.36%</td>
<td>30.93%</td>
</tr>
</tbody>
</table>

103
Table 4-3. Operational conditions during different states at each data collection sites. (Units: flow - veh/h/ln, occ - %, speed - mi/h).

<table>
<thead>
<tr>
<th></th>
<th>10-min before congestion</th>
<th>During incident &amp; congestion</th>
<th>After incident &amp; during congestion</th>
<th>10-min post recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>avg. flow</td>
<td>avg. occ</td>
<td>avg. speed</td>
<td>avg. flow</td>
</tr>
<tr>
<td>Toronto</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No incident</td>
<td>2040</td>
<td>17</td>
<td>60</td>
<td>N/A</td>
</tr>
<tr>
<td>incident before</td>
<td>1799</td>
<td>11</td>
<td>66</td>
<td>985</td>
</tr>
<tr>
<td>incident during</td>
<td>1752</td>
<td>17</td>
<td>56</td>
<td>880</td>
</tr>
<tr>
<td>incident downstream</td>
<td>1407</td>
<td>16</td>
<td>67</td>
<td>946</td>
</tr>
<tr>
<td>Minneapolis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No incident</td>
<td>1881</td>
<td>19</td>
<td>64</td>
<td>N/A</td>
</tr>
<tr>
<td>incident before</td>
<td>1563</td>
<td>14</td>
<td>63</td>
<td>1131</td>
</tr>
<tr>
<td>incident during</td>
<td>1988</td>
<td>17</td>
<td>67</td>
<td>1125</td>
</tr>
<tr>
<td>incident downstream</td>
<td>1650</td>
<td>16</td>
<td>68</td>
<td>1322</td>
</tr>
<tr>
<td>Portland</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No incident</td>
<td>1859</td>
<td>13</td>
<td>55</td>
<td>N/A</td>
</tr>
<tr>
<td>incident before</td>
<td>1394</td>
<td>9</td>
<td>53</td>
<td>889</td>
</tr>
<tr>
<td>incident during</td>
<td>1697</td>
<td>11</td>
<td>54</td>
<td>789</td>
</tr>
<tr>
<td>incident downstream</td>
<td>2061</td>
<td>16</td>
<td>46</td>
<td>528</td>
</tr>
</tbody>
</table>

Table 4-4. Changes in density at the beginning of congestion or incidents at Toronto site.

<table>
<thead>
<tr>
<th></th>
<th>% Change in density at t_{i+1}</th>
<th>% Change in density at t_{i-1}</th>
<th>% Change in density at upstream</th>
<th>% Change in density at downstream</th>
</tr>
</thead>
<tbody>
<tr>
<td>No incident</td>
<td>1-min</td>
<td>8.67%</td>
<td>-17.05%</td>
<td>-35.32%</td>
</tr>
<tr>
<td>(#2)</td>
<td>5-min</td>
<td>35.31%</td>
<td>-10.32%</td>
<td>-29.21%</td>
</tr>
<tr>
<td>Incident</td>
<td>1-min</td>
<td>21.80%</td>
<td>-32.92%</td>
<td>-34.59%</td>
</tr>
<tr>
<td>(#3)</td>
<td>5-min</td>
<td>34.46%</td>
<td>-36.45%</td>
<td>-11.80%</td>
</tr>
</tbody>
</table>
Figure 4-1. Time series speed and flow plots for condition with no incident. A) Time series occupancy, B) Time series speed, and C) Time series flow and speed before, during and after congestion.
Figure 4-2. Time series plots for incidents occurring before breakdown (March 9, 2005).
A) Time series occupancy, B) Time series speed, and C) Time series flow and speed before, during and after congestion.
Figure 4-3. Time series plots for incident during the congested period (Oct 6, 2005). A) Time series occupancy, B) Time series speed, and C) Time series flow and speed before, during and after congestion.
Figure 4-4. Time series plots for incident occurring downstream condition (Sep 19, 2005). A) Time series occupancy, B) Time series speed, and C) Time series flow and speed before, during and after congestion.
Figure 4-5. Time series speed and flow plots for minor incidents (Nov 8, 2005). A) Time series occupancy, B) Time series speed, and C) Time series flow and speed before, during and after congestion.
Figure 4-6. Distributions of the changes in operations for no incident and incident conditions. A) Changes in flow. B) Changes in speed. C) Changes in occ. (Note: T- Toronto, P – Portland site, M – Minneapolis; N- no incident condition, I – incident condition).
Figure 4-7. Density map of Feb 10 2005 without incident
Figure 4-8. Density map of March 9 2005 with incident at 17:17-18:01 at 490DES
Figure 4-9. Density map of May 25 2005 with incident at 9:57-10:27 at 440DES
Figure 4-10. Density map of Oct 6 2005 with incident at 8:02-8:38 at 440DES (5-min)
Figure 4-11. Density map of Oct 6 2005 with incident at 8:02-8:38 at 440DES (1-min)
Figure 4-12. Density Map of Sep 22 2005 with Incident at 17:16-19:46 at 510DES
CHAPTER 5
FREEWAY CAPACITY UNDER INCIDENT CONDITIONS

Freeway capacity is an important metric that is used to evaluate the performance of a facility as well as to manage traffic operations. Freeway capacity under incident conditions is also an important metric, which can be used to optimize system performance in cases of such disruptions. Incidents may block one or more lanes and/or the adjacent shoulder lane. Previous research (Goolsby, M. E., 19711, HCM 2010) has estimated the capacity remaining after an incident based on the number of lanes blocked by the incident. However, the literature has not reported on the relationship between freeway capacity and other incident characteristics (e.g., incident duration), incident type and geometric characteristics. Moreover, the literature has not produced any models relating the capacity before an incident to the capacity after an incident. Such a relationship could help estimate capacity under incident conditions more accurately, and can be used in incident management, queue estimation, etc.

The objectives of this chapter are to:

- Extract and assess the capacity of freeway facilities under non-incident conditions
- Compare values of capacity remaining after incidents to previous values reported in the literature
- Extract and analyze capacity under incident conditions
- Develop models to predict capacity as well as capacity reduction under incident conditions

Data collected from all the five data collection sites are used in this chapter.

**Capacity under Non-incident Conditions**

This section provides the results of the analysis for estimating capacity under non-incident conditions. Data collected from all the five sites (as shown in Table 3-1) were
used. At each site, the range, average, and standard deviation of the four capacity parameters pertaining to non-incident conditions were obtained using the procedure described above and are provided in Table 5-1. The table provides the weighted average (by the number of data points for each site) of each parameter for 2-lane, 3-lane, and 4-lane sites.

Figure 5-1 plots the four capacity parameters by the number of lanes. As shown, all four parameters are the highest for three-lane sections, and the lowest for five-lane section. Statistical analysis results indicate that all the four parameters are significantly higher (at 95% confidence interval) at three-lane freeway than at other freeways. This suggests that three-lane freeways have a higher productivity in terms of vehicles per hour per lane than other freeway facilities. This is different from the HCM 2010, which does not address the relationship between freeway capacity and number of lanes. It is also different from the HCM 2000, which says that the per lane capacity of a freeway segment increases with the number of lanes. Also, breakdown flow is generally higher than the average discharge flow for all the sites. Among the four capacity parameters, breakdown flow has the largest range of values, while average discharge flow generally has the smallest range of values.

It is also observed that the capacities in Table 5-1 are lower than those mentioned in the HCM. For example, the maximum pre-breakdown flow (1-minute peak capacities) is lower than the capacity of a basic freeway segment (15-minute flows) reported in the HCM 2010, which is approximately 2,300 pc/h/ln under a FFS of 60 mi/h under base traffic and geometric conditions (the FFS of the data collection sites are all above 60 mi/h). The average discharge flow is much lower than the queue discharge
flow of 2,000 to 2,300 pc/h/ln, as suggested in HCM 2000. One possible reason of the difference is that the capacity values shown in Table 5-1 are obtained at freeway bottlenecks, which are usually lower than that at basic freeway segment. The standard deviations of the capacity parameters are in the range of 50 to 250 (veh/h/ln).

**Capacity for Incident Conditions**

This section defines capacity for incident conditions, compares incident capacity measured using this database to the results of previous research, and presents models for estimating capacity and capacity reduction during incidents.

**Capacity under Incident Conditions**

Based on the literature review and data analysis, this dissertation will evaluate two parameters in defining incident capacity: the average discharge flow per open lane when both incident and congestion are present, and the minimum 10-minute flow rate when both incident and congestion are present.

Of the five data collections sites, only data collected from three sites (Minneapolis, Portland, and Toronto) are used to analyze capacity for incident conditions, as these were the only ones providing data on the number of lanes affected. The number of lanes affected is not necessarily the number of lanes closed throughout the incident, as this might vary throughout the incident. There are some cases when the number of lanes closed is equal to the total number of lanes of the facility, however, the flow for those cases is not zero. Also, the exact meaning of this variable might be different for different incidents and different sites. At the Portland site, the incident log records ‘the number of lanes closed’, and there is detailed information about the number of lanes closed throughout the incident (the number of lanes closed may change throughout the incident). Thus, at the Portland site, an incident data point might be split into several
data points which would be analyzed separately according to the variation of number of lanes affected throughout the incident.

Apart from the number of lanes affected by the incident, the incident data also include the date, location, start time, duration (except for the Toronto site), and cause of the incident. Some other factors that might affect incident capacity are also collected, including incident category, congestion duration, speed limit, free flow speed, and peak hour factor, etc.

The relationship between parameters of incident capacity and the number of lanes open/affected is illustrated in Table 5-2, based on data analysis of the three sites. The number of lanes open is obtained as the difference between total number of lanes and number of lanes affected. As shown, the average flow per total lanes for incident conditions does not always follow the expected trends. For example, for a facility with no lane open, the average flow is higher when the shoulder is affected than when it is not! This likely occurs because the variable “Number of lanes affected” is not constant throughout the analysis period, and that pattern might also vary from incident to incident. The discrepancies are also likely attributable to sample size. The average flow per open lanes also provides a wide range of results, probably for similar reasons. As shown, its highest value is for ‘1 lane open’ conditions, and it is not much different for ‘2 lanes open’ and ‘3 lanes open’ conditions. The results pertaining to the minimum 10-min flow rate are more reasonable, as this flow increases with the number of lanes open. However, there are some discrepancies with respect to the impact of shoulder closures. Generally, the values of the standard deviations of incident capacities are around 300.
Table 5-3 compares the two parameters of incident capacity at the three sites by the number of lanes open. The table shows that, for all the sites, the average discharge flow per open lane when both incident and congestion are present is the highest for ‘1 lane open’ sites, then for ‘2 lanes open’ sites, and then for ‘0 lane open’ sites. This indicates that only a few vehicles can go through when all the lanes are affected, and that the open lane is used more productively when 1 lane is open than when 2 lanes are open. The minimum 10-min flow rate increases with number of lanes open, except for the ‘2 lanes open’ and ‘1 lane open’ data at the Portland site. The reason might be that the total number of lanes are mostly 2 for ‘1 lane open’ data and are mostly 3 for ‘2 lanes open’ data at the Portland site, thus, the per lane flow at locations with three lanes are lower than that at locations with two lanes. It should be noted that for data along “one lane open” facilities, some of the flows per open lane are higher than expected (above 2000 veh/hr/lane). This might occur, either because there are more than one lane for a portion of the time, or because more than one vehicle can pass through the incident area simultaneously. It is also observed from Table 5-3 that the values of the standard deviations of incident capacities are around 300 and higher than 500 for ‘1 lane open’ condition.

Based on the observations provided above, in this dissertation, the capacity remaining after incidents is calculated in two ways. It is calculated as the ratio of the minimum 10-min flow rate to the average discharge flow for non-incident conditions (averaged for each site), and also as the ratio of the minimum 10-min. flow rate to the 10-min flow before breakdown for non-incident conditions (averaged for each site). For
Comparison purposes, the capacity remaining after incidents under various conditions found by this chapter and previous research are provided in Table 5-4.

Comparing the values shown in Table 5-4 based on the average discharge flow, it is found that:

- For a 2-lane facility, the value of capacity remaining for shoulder affected is close to the values in previous research. However, incident capacity for 1-lane affected and 2-lanes affected conditions are much higher (by about 15%) than that in previous research.

- For a 3-lane facility, the value of capacity remaining for 1-lane affected is close to the values in previous research. However, incident capacity for 2-lanes affected is higher (by about 10%-15%) than that in previous research.

Comparing the values obtained by this chapter based on average flow 10-min before breakdown with previous research, it is found that:

- For a 2-lane facility, the values of capacity remaining is lower by about 10% than previous research for shoulder affected conditions, and are higher by about 10% than previous research for 1-lane affected and 2-lanes affected conditions.

- For a 3-lane facility, the values of percent capacity remaining after incidents are similar to that by Smith et al (2003), but are different from other research by about 10%.

A major reason for the differences in remaining capacity between previous research and this dissertation is probably that previous research (e.g., Smith et al., 2003, Chin et al., 2004) used the peak/maximum of the flow-density curve as capacity under prevailing non-incident conditions, while this dissertation uses the average discharge flow and the average flow 10-min before breakdown for each site as the capacity under non-incident conditions. Thus, the obtained percent of capacity remaining after incidents in this dissertation might be higher. Another reason might be that the percent values obtained in this dissertation are based on the number of lanes affected by incidents. However, in previous research, the remaining capacity was
estimated based on the number of lanes blocked by incidents. The discrepancies might also be due to the fact that incident capacities (mean the minimum 10-min flow rate here) are variables with a large variance. As shown in Table 5-2 and Table 5-3, the standard deviations of the minimum 10-min flow rate are around 300 veh/h/ln, which are larger than that of non-incident capacities (50-250 veh/h/ln, as shown in Table 5-1). Another difference from previous research is that this dissertation analyzes capacity for shoulder plus one lane or two lanes affected conditions.

**Estimate of Capacity/Capacity Reduction under Incident Conditions**

Apart from the number of lanes affected/open, the effects of some other variables on the two parameters of incident capacity are also evaluated, as shown in Table 5-5.

Table 5-5 shows that both parameters of incident capacity are much lower for incidents occurring during congested periods than for incidents occurring before congestion or downstream of the bottleneck. This is because for incidents occurring during congested periods, flows just prior to the incident are much lower than those for other conditions. The minimum 10-min flow is higher at locations with 2 lanes than at locations with 3 lanes. There is no clear pattern of the effects of the other three factors on the two parameters.

It was also observed from scatter plots that, generally, the two parameters of incident capacity decrease slightly with incident duration. However, the trend is small and variation is large. There seems to be no relationship between incident capacity and congestion duration. Thus, these variables are not used in the model development described below.

A regression model was developed to estimate capacity under incident conditions, as well as capacity reduction during incidents.
Based on the above observations, the factors incident duration, incident category, total number of lanes, the number of lanes open, and number of lanes affected, should be further explored in developing incident capacity estimation models. Factors such as incident duration were not found to be related to incident capacity in the data analysis process, and thus are not included in the model. Data with ‘0-lane affected’, ‘shoulder+1-lane affected’, and ‘shoulder+2 -lanes affected’, which are mostly from the Portland site, are not used in the model estimation, as they have few data points (less than 10 points, as shown in Table 5-2). Preliminary analysis with these data points included was not as good. The two parameters of incident capacity were each used as the dependent variable, and it was concluded that the models had a better fit when the minimum 10-min flow rate is used. The estimates of the parameters for minimum 10-min flow rate are shown in Table 5-6.

Using the total capacity reduction (uses average discharge flow as capacity for non-incident conditions and the minimum 10-min flow as capacity for incident conditions), as the dependent variable, the preliminary estimates of the parameters are shown in Table 5-7. In this case, the incident category, total number of lanes, and number of lanes affected are statistically significant at the 95% confidence interval.

Based on the above analysis, the number of lanes open/affected by incidents is found to be the most important factor that influences incident capacity. Another significant factor is incident category. Also, using the minimum 10-min flow rate results in a better model, likely because it matches better the number of lanes affected. The estimated total capacity reduction is illustrated in the following equation,
Conclusions and Recommendations

This chapter studies freeway capacity for both non-incident and incident conditions. The following conclusions are drawn.

Based on the data analyzed in this chapter, three-lane freeways seem to be the most efficient in terms of per lane capacity for non-incident conditions. This finding is based on the observation that three-lane freeways had the highest four parameters of capacity (breakdown flow, etc), compared to two-lane, four-lane and five-lane freeways. This differs from the HCM 2010 which does not address the relationship between freeway capacity and number of lanes, and differs from the HCM 2000 which indicates that per lane capacity increases with the number of lanes. This finding is based on data collected from five freeways, and needs to be further explored with data from additional freeways.

Two parameters were identified for defining incident capacity: average discharge flow per open lane when both incident and congestion are present, and the minimum 10-min flow rate during that same period. It can be concluded that the minimum 10-min flow rate provides a better fit with the data as they are currently reported, because it correlates better with the maximum number of lanes affected.

Apart from the number of lanes affected by incidents, the relationship between incident capacity and other variables was also examined. Multiple linear regression models were developed to estimate incident capacity and capacity reduction under incident conditions, based on parameters such as incident category, number of lanes...
open, and the number of lanes affected. Results show that the model estimating total capacity reduction had a better fit. Three parameters (incident category, total number of lanes and number of lanes affected), were found to be statistically significant at the 95% confidence interval in estimating total capacity reduction. Capacity reduction is higher for incidents occurring during congestion than for incidents occurring before congestion or downstream of the road section, and increases with the number of lanes and the number of lanes affected. There is no linear relationship found between incident capacity / capacity reduction and number of lanes affected.

This dissertation considered capacity for shoulder plus one lane or two lanes affected conditions, however the number of data points was not adequate for those to be included in the model development. Thus it is recommended that additional data be collected for those types of incidents.

With respect to incident data collection, ideally incidents should be reported such that the number of lanes closed throughout the incident is reported. The Portland, Oregon database provides such an incident log, which greatly facilitates the modeling of incident capacity.
<table>
<thead>
<tr>
<th>Location</th>
<th># Lanes</th>
<th>Length (mi)</th>
<th># Data points (days)</th>
<th>Section length (range)</th>
<th>Maximum pre-breakdown flow</th>
<th>Avg. flow for 10 min before breakdown</th>
<th>Avg. discharge flow</th>
<th>Breakdown flow</th>
<th>Maximum pre-breakdown flow</th>
<th>Avg. flow for 10 min before breakdown</th>
<th>Avg. discharge flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minneapolis, MN</td>
<td>2</td>
<td>3</td>
<td>35</td>
<td>1350-2370</td>
<td>1920-2610</td>
<td>1614-2271</td>
<td>1435-1896</td>
<td>1876 (218)*</td>
<td>2181 (163)*</td>
<td>1879 (129)*</td>
<td>1644 (96)*</td>
</tr>
<tr>
<td>Portland, OR</td>
<td>2</td>
<td>7</td>
<td>32</td>
<td>1530-2565</td>
<td>1935-2565</td>
<td>1296-2118</td>
<td>1414-2025</td>
<td>2010 (246)*</td>
<td>2238 (161)*</td>
<td>1858 (151)*</td>
<td>1741 (146)*</td>
</tr>
<tr>
<td><strong>Weighted average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1940 (232)*</td>
<td>2208 (162)*</td>
<td>1869 (140)*</td>
<td>1690 (120)*</td>
</tr>
<tr>
<td>Toronto, Canada</td>
<td>3</td>
<td>6.5</td>
<td>56</td>
<td>1420-2520</td>
<td>1840-2560</td>
<td>1512-2264</td>
<td>1580-2046</td>
<td>2090 (247)*</td>
<td>2330 (162)*</td>
<td>2041 (170)*</td>
<td>1865 (124)*</td>
</tr>
<tr>
<td>Sacramento, CA</td>
<td>3</td>
<td>10.4</td>
<td>35</td>
<td>1440-2280</td>
<td>1860-2460</td>
<td>1597-2154</td>
<td>1183-1843</td>
<td>1943 (199)*</td>
<td>2174 (107)*</td>
<td>1901 (97)*</td>
<td>1563 (142)*</td>
</tr>
<tr>
<td><strong>Weighted average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2033 (229)*</td>
<td>2270 (142)*</td>
<td>1987 (142)*</td>
<td>1749 (131)*</td>
</tr>
<tr>
<td>Sacramento, CA</td>
<td>4</td>
<td>10.4</td>
<td>40</td>
<td>630-2100</td>
<td>1680-2265</td>
<td>948-1962</td>
<td>1100-1756</td>
<td>1750 (256)*</td>
<td>2018 (108)*</td>
<td>1783 (176)*</td>
<td>1567 (115)*</td>
</tr>
<tr>
<td>San Diego, CA</td>
<td>4</td>
<td>2.4</td>
<td>39</td>
<td>1305-2175</td>
<td>1860-2310</td>
<td>1611-1979</td>
<td>1422-1890</td>
<td>1868 (160)*</td>
<td>2075 (113)*</td>
<td>1829 (86)*</td>
<td>1665 (85)*</td>
</tr>
<tr>
<td><strong>Weighted average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1808 (209)*</td>
<td>2046 (110)*</td>
<td>1806 (131)*</td>
<td>1615 (100)*</td>
</tr>
<tr>
<td>San Diego, CA</td>
<td>5</td>
<td>2.4</td>
<td>34</td>
<td>1392-2076</td>
<td>1812-2076</td>
<td>1660-1866</td>
<td>1473-1741</td>
<td>1774 (160)*</td>
<td>1928 (70)*</td>
<td>1756 (58)*</td>
<td>1635 (66)*</td>
</tr>
</tbody>
</table>

* Numbers in the parentheses indicate the standard deviations (S.D.) of the flow parameters.
Table 5-2. Incident capacity and number of lanes affected

<table>
<thead>
<tr>
<th># Lanes open</th>
<th># Lanes affected</th>
<th># Data points (days)</th>
<th>Avg. flow per total lanes (veh/h/ln)</th>
<th>Avg. flow per open lanes (veh/h/open lane)</th>
<th>Minimum 10-min flow rate (veh/h/ln)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>11</td>
<td>418</td>
<td>836*</td>
<td>237</td>
</tr>
<tr>
<td>0</td>
<td>shoulder +2</td>
<td>6</td>
<td>763</td>
<td>1289*</td>
<td>410</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>28</td>
<td>1004</td>
<td>2031</td>
<td>858</td>
</tr>
<tr>
<td>1</td>
<td>shoulder +1</td>
<td>2</td>
<td>798</td>
<td>1595</td>
<td>443</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>8</td>
<td>756</td>
<td>2267</td>
<td>541</td>
</tr>
<tr>
<td>2</td>
<td>Shoulder</td>
<td>14</td>
<td>1383</td>
<td>1383</td>
<td>1268</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>28</td>
<td>1390</td>
<td>1390</td>
<td>792</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1388</td>
<td>1388</td>
<td>1180</td>
</tr>
</tbody>
</table>

Note: *assumes there is a passage for the vehicles on the shoulder or between the lanes, or that all lanes are closed only for a brief amount of time.

Table 5-3. Comparison of incident capacity at different sites

<table>
<thead>
<tr>
<th># Lanes open</th>
<th>Minnesota site</th>
<th>Portland site</th>
<th>Toronto site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. discharge flow during incident &amp; congestion (veh/h/open lane)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>1043*</td>
<td>560</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>1592</td>
<td>487</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>1428</td>
<td>264</td>
</tr>
<tr>
<td>3</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Minimum 10-min flow rate (veh/h/ln)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>305</td>
<td>159</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>622</td>
<td>309</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>1309</td>
<td>260</td>
</tr>
<tr>
<td>3</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Note: *assumes there is one pass for the vehicles on the shoulder or between the lanes.
Table 5-4. Comparison of percent of freeway capacity available under incident conditions

<table>
<thead>
<tr>
<th>Author</th>
<th>Number of lanes</th>
<th>Lanes blocked</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>shoulder</td>
<td>1 lane</td>
<td>2 lanes</td>
<td>3 lanes</td>
</tr>
<tr>
<td>Goolsby, 1971</td>
<td>3 (27 data points)</td>
<td>0.67</td>
<td>0.50</td>
<td>0.21</td>
<td>0.00</td>
</tr>
<tr>
<td>HCM 2010</td>
<td>2</td>
<td>0.81</td>
<td>0.35</td>
<td>0.00</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.83</td>
<td>0.49</td>
<td>0.17</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.85</td>
<td>0.58</td>
<td>0.25</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.87</td>
<td>0.65</td>
<td>0.40</td>
<td>0.20</td>
</tr>
<tr>
<td>Smith et al., 2003</td>
<td>3 (27 data points)</td>
<td>N/A</td>
<td>0.37</td>
<td>0.23</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.75</td>
<td>0.32</td>
<td>0.00</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.84</td>
<td>0.53</td>
<td>0.22</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.89</td>
<td>0.56</td>
<td>0.34</td>
<td>0.15*</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.93*</td>
<td>0.75</td>
<td>0.50</td>
<td>0.20*</td>
</tr>
<tr>
<td>Lu and Elefteriadou, 2011 (average discharge flow)</td>
<td>2 (60 data points)</td>
<td>0.77</td>
<td>0.50</td>
<td>0.14</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>3 (30 data points)</td>
<td>N/A</td>
<td>0.43</td>
<td>0.32</td>
<td>N/A</td>
</tr>
<tr>
<td>Lu and Elefteriadou, 2011 (avg. flow 10-min before breakdown)</td>
<td>2 (60 data points)</td>
<td>0.68</td>
<td>0.46</td>
<td>0.13</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>3 (30 data points)</td>
<td>N/A</td>
<td>0.40</td>
<td>0.29</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*Note: * assumed.
Table 5-5. Effects of various factors on parameters of incident capacity

<table>
<thead>
<tr>
<th>Factors</th>
<th>Incident category</th>
<th>Incident location</th>
<th>Total # lanes</th>
<th>Speed limit (mph)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Number of data point (days)</td>
<td>31</td>
<td>69</td>
<td>49</td>
<td>51</td>
</tr>
<tr>
<td>Avg. discharge flow during inci. &amp; cong. (Veh/h/open lane)</td>
<td>1481</td>
<td>1616</td>
<td>1632</td>
<td>1519</td>
</tr>
<tr>
<td>Minimum 10-min flow rate (veh/h/ln)</td>
<td>658</td>
<td>819</td>
<td>728</td>
<td>809</td>
</tr>
</tbody>
</table>

0-lane open

| Number of data point (days)          | 6                | 11                | 11            | 6                | 17          | 0           | 15          | 2           |
| Avg. discharge flow during inci. & cong. (Veh/h/open lane) | 978              | 1005              | 1077          | 845              | 996         | N/A         | 989         | 1043        |
| Minimum 10-min flow rate (veh/h/ln)  | 299              | 297               | 353           | 198              | 298         | N/A         | 297         | 305         |

1-lane open

| Number of data point (days)          | 13               | 25                | 23            | 15               | 29          | 9           | 25          | 13          |
| Avg. discharge flow during inci. & cong. (Veh/h/open lane) | 1890             | 2145              | 2068          | 2043             | 2004        | 2232        | 2077        | 2021        |
| Minimum 10-min flow rate (veh/h/ln)  | 694              | 809               | 810           | 708              | 841         | 540         | 856         | 604         |

2-lane open

| Number of data point (days)          | 12               | 32                | 14            | 30               | 15          | 29          | 8           | 36          |
| Avg. discharge flow during inci. & cong. (Veh/h/open lane) | 1290             | 1419              | 1369          | 1392             | 1390        | 1382        | 1275        | 1409        |
| Minimum 10-min flow rate (veh/h/ln)  | 799              | 995               | 856           | 982              | 1231        | 805         | 765         | 986         |

Note: Incident location B – bottleneck, NB - non-bottleneck.
Incident category 1- incidents occurring during congestion, 0 – incidents before congestion or downstream.

Table 5-6. Estimates of parameters for regression to estimate the minimum 10-min flow rate (veh/h/ln)

| Parameter | Description                      | Estimate | Error  | t Value | Pr>|t| |
|-----------|----------------------------------|----------|--------|---------|------|
| inccat    | Incident category:               |          |        |         |      |
|           | 1 - during congestion           | -165.8   | 67.9   | -2.44   | 0.0166|
| nlaneaff  | One lane affected                | 885.5    | 46.2   | 19.17   | <.0001|
| nlaneaff  | Two lanes affected               | 408.5    | 69.6   | 5.87    | <.0001|
| nlaneaff  | Shoulder affected                | 1291.2   | 79.0   | 16.35   | <.0001|
| R-square  |                                  | 0.49     |        |         |      |
| Root MSE  |                                  | 293.20   |        |         |      |
Table 5-7. Estimates of parameters for regression to estimate the total capacity reduction (veh/h)

| Parameter  | Description           | Estimate | Error  | t value | Pr>|t| |
|------------|-----------------------|----------|--------|---------|------|
| inccat     | Incident category:    |          |        |         |      |
|            | 1-during congestion   | 386.5    | 167.4  | 2.31    | 0.0234|
| nlane      | Total number of lanes | 1361.1   | 166.6  | 8.17    | <.0001|
| nlaneaff   | One lane affected     | -1171.6  | 442.7  | -2.65   | 0.0097|
|            | Two lanes affected    | -76.7    | 442.6  | -0.17   | 0.8629|
|            | Shoulder affected     | -2008.3  | 387.7  | -5.18   | <.0001|
| R-square   |                       | 0.67     |        |         |      |
| Root MSE   |                       | 718.40   |        |         |      |
Figure 5-1. Capacity parameters by number of lanes
Previous research has investigated the occurrence of flow breakdown on freeways, which is defined as the beginning of congestion. It represents the transition from relatively free-following traffic to congestion, or stop-and-go traffic. There are several models reported in the literature that predict the occurrence of breakdown (Elefteriadou et al., 1995, Evans et al., 2001, Brilon et al., 2005). Generally, these models predict breakdown due to demand (demand-induced breakdown) as a function of flow: the higher the freeway and ramp flows, the higher the probability of breakdown.

One of the most popular methods used to develop breakdown models is the Product Limit Method (PLM) that is based on the work by Brilon et al (2005). The method is also called the Kaplan-Meier (1958) method. This method uses lifetime data analysis for estimating the time until failure of mechanical parts or the duration of human life. Generally, the probability of failure increases with time, which is true in the relationship between breakdown occurrence and flow.

Even though this method has been used to describe the probability of demand-induced breakdown, there is limited research on the phenomenon of freeway breakdown induced by incidents (incident-induced breakdown). The same methods may be used to examine the relationship between incident-induced breakdown and flow. This is because the mechanisms of demand-induced breakdown and incident-induced breakdown have certain similarities. Microscopic-level driver behaviors, such as car-following and lane changing, can lead to flow breakdown, and can also lead to incidents. Therefore, it is possible that methods that predict demand-induced
breakdown can also be used in predicting the probability of incident-induced breakdown.

The objective of the chapter is to apply the product limit method to estimate the occurrence of breakdown under both non-incident conditions and incident conditions, and to compare the models developed for different conditions and different sites. Several different models are developed based on various traffic flow parameters (flow, occupancy, difference between speed limit and operating speed, standard deviation of speed, and average variation of speed) and the results are compared to determine which parameters are most useful. The types of models developed in this chapter could ultimately be used in formulating and applying real-time advanced traffic management systems and various freeway management strategies.

**Overview of the Product Limit Method (PLM)**

The PLM is based on the work by Kaplan and Meier (1958), which uses lifetime data analysis for estimating the time until failure of mechanical parts or the duration of human life. It takes advantage of the fact that the N-year survival rate is equal to the product of all of the survival rates of the individual intervals leading up to time N. The lifetime distribution function is given by:

\[
F(t) = 1 - S(t) \quad \text{(Eq. 6-1)}
\]

Where,

\[
F(t) = P(T \leq t) \quad \text{is the distribution function of a lifetime, } T \text{ is lifetime duration}
\]

\[
S(t) = P(T > t) \quad \text{is the survival function}
\]

The product-limit estimator of the survival function is given by:
\[ \hat{S}(t) = \prod_{t_j \leq t} \frac{n_j - \delta_j}{n_j} \]  
(Eq. 6-2)

Where,

- \( n_j \) is the number of items with lifetime \( T \geq t_j \)
- \( \delta_j \) is the number of failures or deaths at time \( t_j \)

In the context of freeway breakdown, the event of a failure at time \( t \) is the event of a breakdown at volume \( q \), and the lifetime is the maximum pre-breakdown volume.

Thus, equation 6-1 can be modified to calculate the distribution of breakdown probability, as shown in the following relationship:

\[ F(q) = P(q_i \leq q) = 1 - P(q_i > q) \]  
(Eq. 6-3)

Where,

- \( F(q) \) is breakdown probability distribution
- \( q \) is the observed traffic volume (veh/h/ln)
- \( q_i \) is the traffic volume in interval \( i \), which is the one prior to the drop in speeds, defined as the breakdown flow (veh/h/ln)
- \( P(q_i > q) \) is the probability that the breakdown volume is greater than the observed volume (i.e., the probability that no breakdown will occur up to that volume)

Equation 6-2 can then be written as follows:

\[ \hat{p}(q_i > q) = \hat{S}(q) = \prod_{i, q_i < q} \frac{k_i - d_i}{k_i}, i \in B \]  
(Eq. 6-4)

Where,

- \( q \) is the observed traffic volume (veh/h/ln)
- \( q_i \) is the traffic volume in interval \( i \), which is the one prior to the drop in speeds, defined as the breakdown flow (veh/h/ln)
$k_i$ is the number of intervals with a traffic volume of $q \geq q_i$

d_i$ is the number of breakdowns at a volume of $q_i$

$\{B\}$ is the set of breakdown intervals

The corresponding breakdown probability function is:

$$F(q) = 1 - \prod_{i|q_i \geq q} \frac{k_i - d_i}{k_i}, i \in \{B\}$$

(Eq. 6-5)

If each observed volume that causes breakdown is considered separately (that is, only one observation of breakdown for every maximum pre-breakdown volume, $q_i$, $d_i = 1$), then equation 6-5 becomes:

$$F(q) = 1 - \prod_{i|q_i \geq q} \frac{k_i - 1}{k_i}, i \in \{B\}$$

(Eq. 6-6)

In applying the PLM, 1-minute time intervals are used to identify flow breakdown. Brilon et al. (2005) pointed out that only rather short observation intervals (ideally one minute or even less) are useful for analysis, otherwise the causal relationship between traffic volume and breakdown would be too weak. A total of 10 minute intervals before the demand-induced breakdown are used: the last 1-min is used to represent breakdown flow in the equations and the others are non-breakdown flows. This determination is based on the work of Elefteriadou et al. (2009). Data analysis showed that flow rates much earlier than the breakdown do not contribute to the calculation results.

As mentioned earlier, the mechanisms of breakdown and incidents have certain similarities and thus the methods that predict the probability of for demand-induced breakdown could also be used in predicting the probability of incident-induced breakdown.
Application of the Product-Limit Method for Incident Conditions

In order to apply the PLM to model incident-induced breakdown, the equations and respective parameters are translated to consider parameters related to incidents, rather than those related to breakdown.

The incident-induced breakdown interval is the one immediately preceding an incident which has led to congestion. Several models are developed based on several different traffic parameters. Flow and occupancy, which have been used to predict the probability of breakdown, are used, for comparison purposes. Models are also developed based on other parameters that might be related to breakdown occurrence: speed difference (speed limit minus speed), 5-min standard deviation of speed, and 5-min average variation of speed (CVS).

Based on the above, equation 6-3 becomes:

\[ F(q_{inc}) = P(q_{inc} \leq q) = 1 - P(q_{inc} > q) \]

(Eq. 6-7)

Where,

- \( F(q_{inc}) \) is the incident-induced breakdown probability distribution
- \( P(q_{inc} > q) \) is the probability that the incident-induced breakdown volume is greater than the observed volume (i.e., the probability that no incident-induced breakdown will occur up to that volume). According to the PLM it is given by:

\[ \hat{P}(q_{inc} > q) = \prod_{i : q_{inc} \leq q} \frac{k_{i} - inc_{i}}{k_{i}}, i \in \{B\} \]

(Eq. 6-8)

The corresponding incident-induced breakdown probability function is given by:

\[ F(q_{inc}) = 1 - \prod_{i : q_{inc} \leq q} \frac{k_{i} - inc_{i}}{k_{i}}, i \in \{B\} \]

(Eq. 6-9)

Where,
\( q \) is the observed traffic volume (veh/h/ln)

\( q_{inci} \) is traffic volume in interval \( i \), which is the one prior to the incident-induced breakdown (veh/h/ln)

\( k_i \) is the number of intervals with a traffic volume of \( q \geq q_{inci} \)

\( inc_i \) is the number of incident-induced breakdowns at a volume of \( q_{inci} \)

\( \{B\} \) is the set of incident-induced breakdown intervals (1-minute observations)

\( \{F\} \) is the set of intervals prior to the incident-induced breakdown

If each observed volume that causes an incident is considered separately (that is, only one observation of incident-induced breakdown for every maximum pre-incident volume, \( inc_i = 1 \)), then equation 6-9 becomes:

\[
F(q_{inc}) = 1 - \prod_{i:q_{inci} \leq q} \frac{k_i - 1}{k_i}, i \in \{B\}
\]

(Eq. 6-10)

Similarly to the analysis for demand-induced breakdown, 1-minute time intervals are used in the analysis, and data for 10 minutes before the incident-induced breakdown were used.

The next three sections present the use of the product limit method to estimate the occurrence of breakdown both due to demand and due to incidents, and makes the comparisons.

**Probability of Demand-induced Breakdown**

This section first presents the data used and describes in some detail the estimation of the probability of demand-induced breakdown. It then presents the results based on the five traffic parameters identified in the previous section and based on these results recommends the parameter that is most appropriate to estimate the probability of demand-induced breakdown.
Description of Data and Analysis Procedure

Data collected from three sites are used in this chapter: Toronto, Minneapolis, and Portland, as they have more incidents verified. Table 6-1 provides an overview of the study site characteristics for the sites used in this chapter, and the data period.

Data during non-incident conditions (with no incident) were analyzed at the following bottleneck locations: two bottlenecks at the Toronto site, one bottleneck at the Minneapolis site, and four bottlenecks at the Portland site (two locations during the AM period and two locations during the PM period). At the Toronto site, only data at the main bottleneck 510 DES (which has much more breakdown data points than the other bottleneck) are used in developing the demand-induced breakdown models.

These data are used in developing a probabilistic model of demand-induced breakdown, based on each of the five traffic parameters. Models for each site and for all the data are developed. The following steps are followed:

(1) For each demand-induced breakdown occurrence, obtain the average 1-minute parameter for 10 minutes prior to the breakdown. The parameters 5-minute standard deviation of speed (5-min std.v) and 5-minute average variation of speed (5-min cvs) are calculated for each interval based on the speeds during the previous 5 minutes. Thus, speeds from the 15 minutes prior to the breakdown are used in calculating these two parameters. Use the parameter of the last minute as the breakdown interval and the other values as the non-breakdown intervals.

(2) Calculate the breakdown probability using equations 6-3 through 6-6.

(3) Develop the probability distribution curve as a function of the subject parameter.

Demand-induced Breakdown Models

Flow-based model

Flow has been the parameter mostly used in developing such models in previous research. The results for the set of data used in this chapter are shown in Figure 6-1. The horizontal axis shows the flow rate, while the vertical axis indicates the probability
of demand-induced breakdown. The figure shows that the probability of demand-induced breakdown is very small for flows less than 1,600 veh/h/ln, and increases to about 40% when the flow reaches 2,500 veh/h/ln. When flow rate exceeds 2,500 veh/h/ln, the probability of demand-induced breakdown increases greatly.

As shown in Figure 6-1, the probability curves at the three sites do not differ much when flow is low (below 2,100 veh/h/ln). The probability values are in the range of 5% to 11% at the three sites for flow rates around 2,000 veh/h/ln. However, the differences in probabilities at the three sites increase when flow increases (mostly when it exceeds 2,200 veh/h/ln). More specifically, the Toronto site has a lower probability of demand-induced breakdown than the other two sites at the same flow value. For example, at a flow rate 2,220 veh/h/ln, the probability is about 12% at the Toronto site and about 27%, and 23% at the Minneapolis and Portland sites, respectively. This indicates that breakdown is triggered at a lower per lane flow rate at the other two sites than at the Toronto site. The reason might be that there are a total of three lanes at the Toronto site, while only two lanes at the Minneapolis and Portland sites (see Table 6-1). This finding is consistent with an earlier research paper, which found that three-lane facilities are more productive (higher throughput per lane) than two-lane or four lane facilities (Lu and Elefteriadou, 2011).

**Occupancy-based model**

Next, the probability of demand-induced breakdown is estimated based on occupancy, as shown in Figure 6-2.

Figure 6-2 shows that when the occupancy is below 13%, probabilities of demand-induced breakdown are close to zero and are nearly identical at the three sites. However, there are large differences in the models for the three sites when occupancy
exceeds 15%: the probability of demand-induced breakdown at the Portland site increases greatly when occupancy exceeds 15%, while at the other two sites it starts to increase significantly when the occupancy exceeds 20%. Thus, the probability of demand-induced breakdown at the Portland site is much higher than at the other two sites at the same occupancy. One possible reason is that the speed limit at the Portland site is lower (55 mph) than at the other two sites (60 mph). Thus, the same (breakdown) flow occurs at a lower speed and higher occupancy. Another, potentially related reason, is that there are comparatively more on/off-ramps (especially off-ramps) per mile at the Portland site than at the other two sites (see Table 6-1). At the Toronto site, all the 52 data points are within the main bottleneck and there is no off-ramp downstream. At the Minneapolis site, only 2 of the total 34 data points have off-ramps immediately downstream. However, at the Portland site, 32 of the total 33 data points represent locations with off-ramps immediately downstream within 2,500 feet. Thus it is speculated that the lower speed limit, coupled with the presence of a downstream off-ramp results in higher occupancies.

To investigate whether the presence of downstream off-ramps affected the probability of demand-induced breakdown, a comparison of probabilities of demand-induced breakdown was conducted for the two different bottleneck locations at the Toronto site. As indicated earlier, the Toronto site has two bottlenecks: one main bottleneck at detector 510 DES which has 52 non-incident data points (as shown in Figure 6-1 and Figure 6-2) and is free from downstream flow effects, and one at detector 480 DES which has 9 non-incident data points with an off-ramp immediately downstream within about 2000 ft. The second set of data was not used in the previous
The speed limits at the two bottlenecks are the same. The probabilities of demand-induced breakdown at the two different bottleneck locations based on flow and occupancy are shown in Figure 6-3 A) and B). As shown, the differences in probability of demand-induced breakdown at the two bottleneck locations are very small when estimated based on flow. However, the differences are very large when the models are based on occupancy: the probability of demand-induced breakdown at the bottleneck 480 DES, that is, the bottleneck with an off-ramp immediately downstream, is much higher than that at the bottleneck 510 DES at the same occupancy, and increases greatly when occupancy exceeds 15%. This observation is consistent with the analysis of Figure 6-2, and indicates that traffic at bottlenecks with off-ramps immediately downstream would break down at a lower occupancy than at bottlenecks without off-ramps immediately downstream. It also indicates that at bottlenecks with off-ramps immediately downstream, breakdown occurrence is more sensitive to occupancy than to flow, and thus occupancy might be a better indicator of breakdown for this type of design.

**Speed difference-based model**

The probability of demand-induced breakdown is estimated based on the difference between speed limit (55 mph for Portland site and 60 mph for the other two sites) and operating speed. To adjust for the differences in speed limits at different sites, the speed difference is normalized by dividing by the speed limit. The estimated probability of demand-induced breakdown is shown in Figure 6-4. The horizontal axis indicates the percent of speed lower than the speed limit, while the vertical axis shows the probability of demand-induced breakdown.
As shown, the relationship for speeds above the speed limit is very similar between the three sites. For speeds lower than the speed limit however there are noticeable differences, and those increase as the operating speeds drop. The Toronto site has a lower probability of demand-induced breakdown than the other two sites. For example, at speed 16% lower than the speed limit, the probability of demand-induced breakdown is about 20% at the Toronto site and about 29% and 37% at the Minneapolis and Portland sites, respectively. This indicates that breakdown is triggered at a higher relative speed at the other two sites than at the Toronto site. Similar to the results based on flow, the reason for these differences could be that there are a total of three lanes at the Toronto site and two lanes at the Minneapolis and Portland site.

5-min standard deviation of speed-based model

The probability of demand-induced breakdown is also estimated based on the 5-minute standard deviation of speed (5-min std.v). The probability distribution is shown in Figure 6-5.

As shown, the differences in probabilities at the three sites are very large when the standard deviation of speed exceeds approximately 6.5. Below that value, the probabilities at the Minneapolis site are lower than at the other two sites. The distribution curves do not show any clear pattern of differences between the three sites.

5-min average variation of speed-based model

The probability of demand-induced breakdown is estimated based on the 5-minute average variation of speed (5-min cvs). The probability distribution is shown in Figure 6-6. As shown, the probability of demand-induced breakdown increases with the average variation of speed.
Similar to the models based on the standard deviation of speed, the differences in probabilities at the three sites are very large when the average variation of speed exceeds approximately 0.1. Below that value, the differences between the three sites are small. Generally, the distribution curves do not show any clear pattern of differences between the three sites.

In summary, the estimated demand-induced breakdown models seem to be different at the three sites as a function of the number of lanes: there are a total of three lanes at the Toronto site and two lanes at the Minneapolis and Portland sites. This is consistent with previous research which has speculated that three-lane facilities are more productive in terms of per-lane capacity than two-lane facilities. The curves also seem to be different based on the presence of downstream off-ramps: the Portland site has comparatively more on/off-ramps per mile and has off-ramps immediately downstream of the breakdown locations, thus demand-induced breakdown occurs at a lower occupancy at the Portland site than at the other two sites. Comparably, the flow-based models are the most consistent at the three sites, but they do not show some of the differences between sites that can be seen using occupancy. The estimated models based on the 5-min standard deviation of speed and average variation of speed, generally, do not show significant differences between the three sites, but they are presented here for purposes of comparison to the incident-induced breakdown models.

**Probability of Incident-induced Breakdown**

This section first discusses the process for estimating the probability of incident-induced breakdown. It then presents the results based on the five traffic parameters identified above.
Process for Investigating the Probability of Incident-induced Breakdown

Only incidents that occur immediately before congestion are used in this dissertation. Incident-induced breakdown probability models are developed based on the five parameters used in developing the demand-induced breakdown models. The following steps are followed:

1. For each incident-induced breakdown occurrence, obtain the average 1-minute parameter for 10 minutes prior to the congestion. The parameters 5-minute standard deviation of speed (5-min std.v) and 5-minute average variation of speed (5-min cvs) are calculated for each interval based on the speeds during the previous 5 minutes. Thus, speeds from the 15 minutes prior to the incident-induced breakdown are used in calculating these two parameters. Use the parameter of the last minute before breakdown as the incident-induced breakdown flow.

2. Obtain the same amount of data points with incidents that have occurred during similar time periods but have not caused congestion.

3. Calculate the incident-induced breakdown probability using equations 6-7 through 6-10.

4. Develop the probability distribution curve as a function of the subject parameter.

The results of the analysis are presented in the next subsection.

Incident-induced Breakdown Models

The resulting incident-induced breakdown models are shown in Figure 6-7. As shown in Figure 6-7 A), when using flow in developing the models, there are no major differences in the probabilities of incident-induced breakdown at the three sites. When flow is below approximately 1,600 veh/h/ln, the relationship is fairly flat and the differences at the three sites are very small. In the range of 1,600 to 2,000 veh/h/ln, the slope increases and the differences between the three sites increase somewhat. For example, at flow 1,900 veh/h/ln, the probability of incident-induced breakdown is about 9%, 15% and 8% at the Minneapolis, Portland and Toronto sites, respectively.
When flow exceeds 2,000 veh/h/ln, there are few data points and the probabilities of incident-induced breakdown increase greatly. The Portland site shows a higher probability of incident-induced breakdown than the other two sites. This may be due to the scarcity of data points at this range, as breakdown has likely already occurred due to high demand. Compared to the probability of demand-induced breakdown (Figure 6-1), the probability of incident-induced breakdown is higher at the same flow. For example, for the Portland site and for a flow of 2,000 veh/h/ln, the probability of incident-induced breakdown is about 18% and the probability of demand-induced breakdown is about 10%.

When using occupancy in developing the models (Figure 6-7 B)), the differences among the three sites are larger: the probabilities of incident-induced breakdown are the largest at the Portland site and the smallest at the Minneapolis site. Similarly to the differences observed in modeling demand-induced breakdown, the lower speed limit of the Portland site, coupled with the presence of a downstream off-ramp likely results in increased occupancy. Compared to the probability of demand-induced breakdown (Figure 6-2), the probability of incident-induced breakdown is higher at the same occupancy. For example, for the Portland site and for an occupancy of 15% the probability of incident-induced breakdown is about 15% and the probability of demand-induced breakdown is about 9%.

Using the other three speed parameters in developing the models (Figure 6-7 C) to Figure 6-7 E)), the patterns are very similar to those shown for demand-induced breakdown. As the speed parameters increase, the differences among the three sites increase. Compared to the probability of demand-induced breakdown (Figures 6-4 to
Figure 6-6), the probability of incident-induced breakdown is lower at the same speed parameters. For example, for the Portland site and for a standard deviation of 6, the probability of incident-induced breakdown is about 15% and the probability of demand-induced breakdown is about 28%.

In summary, the developed probabilities of incident-induced breakdown have similar shape to the probability of demand-induced breakdown based on the same parameter. Similar to the non-incident conditions, the differences of incident-induced breakdown probabilities among the three sites are the smallest when the model is developed based on flow. The incident-induced breakdown models based on the three speed parameters have similar patterns to those for demand-induced breakdown.

**Comparison of Demand-induced and Incident-induced Breakdown Models**

This section compares the models for demand-induced breakdown and incident-induced breakdown. The comparisons based on the five parameters selected are presented in Figure 6-8. Results are shown only for all sites grouped, as the comparisons at each site show very similar patterns.

As shown in Figure 6-8 A), the probability of incident-induced breakdown is higher than the probability of demand-induced breakdown when the flow is between approximately 1,200 and 2,200 veh/h/ln. For example, at a flow rate of 1,800 veh/h/ln, the probability of incident-induced breakdown is about 8%, and the probability of demand-induced breakdown is about 3%. Above 2,000 veh/h/ln, there are very few data points for incident conditions, and the probability of demand-induced breakdown increases much more. This indicates that for the middle range of flows, congestion is more likely to occur as a result of an incident than due to demand. Also, for incident conditions, the ranges of flow and occupancy at which congestion may result are lower
than the respective range for demand-induced breakdown. For example, congestion results when flows are in the range of 210 to 2,295 veh/h/ln for incident conditions, and in the range of 1,350 to 2,565 veh/h/ln for non-incident conditions.

The comparisons based on occupancy (Figure 6-8 B)) show similar patterns to the comparisons based on flow. However, for the other three speed parameters (Figures 6-8 C) – E)), the comparison results are quite different: the probability of demand-induced breakdown is higher than the probability of incident-induced breakdown at the same speed parameter, and the differences increase with the parameters. This indicates that, for incident conditions, speed does not drop as dramatically prior to breakdown, and speed variability is not as high as prior to a demand-induced breakdown. Conversely, high speed variability is more likely to result in a demand-induced breakdown than an incident. For example, when the 5-min standard deviation is 8, there is a probability of 0.4 for a demand-induced breakdown and a probability of 0.12 for an incident-induced breakdown.

Another important observation is that the upper limit of these speed related parameters is higher for incident-induced breakdown than the range observed for demand-induced breakdown. For example, for incidents congestion was observed to start when the 5-min standard deviation of speed reached 13.7, while that upper limit for non-incident conditions is 10.5. Similar differences can be observed for the other two speed parameters.

**Conclusions**

The PLM has been used in previous research to estimate the probability of demand-induced breakdown. This chapter uses this method to estimate the probability of demand-induced breakdown, as well as incident-induced breakdown and compare
the two sets of models. Data collected from three freeways in North America were used in developing the models, based on five traffic parameters: flow, occupancy, the difference between speed limit and speed, 5-min standard deviation of speed, and 5-min average variation of speed. The following conclusions are drawn:

(1) The probability of demand-induced breakdown curves as a function of flow are similar among the three sites. However the two-lane sites have a higher probability of breakdown than a three lane site at the same per lane flow. This is consistent with previous research which has reported that three-lane facilities are more productive in terms of per-lane capacity than two-lane facilities or four-lane facilities.

(2) The probability of demand-induced breakdown curves as a function of occupancy show some interesting differences between sites. Bottlenecks with off-ramps immediately downstream were found to break down at a lower occupancy than bottlenecks without off-ramps downstream. In this case breakdown occurrence is more sensitive to occupancy than to flow, and thus occupancy might be a better indicator of breakdown for this type of design.

(3) The incident-induced breakdown curves can be developed similarly to the demand-induced breakdown curves. The patterns of the curves are similar to the demand-induced breakdown curves in that flow-based curves do not show significant differences between sites, but occupancy-based curves do.

(4) The distributions of demand-induced breakdown and incident-induced breakdown are compared based on each of the five parameters. For the middle range of flows (below 2,000 veh/h/ln), congestion is more likely to occur as a result of an incident than due to demand. When flow exceeds 2,000 veh/h/ln, there are few data points with incident-induced breakdown. Incident-induced breakdown generally occurs at a lower range of flow and occupancy than the demand-induced breakdown.

(5) For incident conditions, speed does not drop as dramatically prior to breakdown, and speed variability is not as high as prior to a demand-induced breakdown. Conversely, high speed variability is more likely to result in a demand-induced breakdown than an incident. Also, the upper limit of the speed related parameters is higher for incident-induced breakdown than the range observed for demand-induced breakdown.

The differences observed in the breakdown probability models between demand-induced and incident-induced breakdown could potentially be used in detecting congestion and incidents and in differentiating between the two events. As shown,
there are some important differences in the range of speed, flow and occupancy at which each event occurs. Furthermore, there are differences in the probability of occurrence of each of these two events. Additional research is required however to be able to use this information in a manner that would be useful to practitioners and agencies.

The types of curves developed in this chapter could also be very helpful in traffic management of freeway facilities. As shown, there are distinct patterns in the probability of breakdown of different sites, and these should be further explored to identify what design elements result in reduced probability of breakdown (both demand-induced, and incident-induced). Furthermore, such curves could be developed for other sites using the method described here, and used when implementing ramp metering or Variable Speed Limit algorithms to optimize freeway operations. Potential such applications are explored in Elefteriadou et al. (2009).

Finally, it is recommended to extend the PLM for predicting the occurrence of incidents, as the mechanisms of breakdown and incidents have certain similarities.
Table 6-1. Summary of site characteristics and data available

<table>
<thead>
<tr>
<th>Site</th>
<th>Location</th>
<th>Length (mi)</th>
<th>Total # of lanes</th>
<th>Speed limit (mph)</th>
<th># of on/off ramps</th>
<th>Data period</th>
</tr>
</thead>
<tbody>
<tr>
<td>QEW</td>
<td>Toronto, Canada</td>
<td>6.5</td>
<td>3</td>
<td>60</td>
<td>8 / 3</td>
<td>01/2005-12/2005</td>
</tr>
<tr>
<td>I-494 SB</td>
<td>Minneapolis, MN</td>
<td>3</td>
<td>2</td>
<td>60</td>
<td>2 / 2</td>
<td>09/2006-08/2007</td>
</tr>
<tr>
<td>OR 217 SB</td>
<td>Portland, OR</td>
<td>7</td>
<td>2</td>
<td>55</td>
<td>8 / 9</td>
<td>11-12/2007</td>
</tr>
</tbody>
</table>
Figure 6-1. Probability of demand-induced breakdown based on flow

Figure 6-2. Probability of demand-induced breakdown based on occupancy
Figure 6-3. Probability of demand-induced breakdown at the two bottleneck locations at Toronto site. A) Probability based on flow, B) Probability based on occupancy.
Figure 6-4. Probability of demand-induced breakdown based on normalized speed difference

Figure 6-5. Probability of demand-induced breakdown based on 5-min std.v
Figure 6-6. Probability of demand-induced breakdown based on 5-min cvs
Figure 6-7. Probability of incident-induced breakdown. A) Flow-based model, B) Occupancy-based model, C) Normalized speed difference-based model, D) 5-min std. v-based model, and E) 5-min cvs-based model.
Figure 6-8. Comparison of probabilities of demand-induced breakdown and incident-induced breakdown. A) Flow-based model, B) Occupancy-based model, C) Normalized speed difference-based model, D) 5-min std.v-based model, and E) 5-min cvs-based model.
CHAPTER 7
FREEWAY INCIDENT DETECTION USING LIKELIHOOD OF INCIDENT FUNCTIONS

In previous chapter, it is recommended to extend the Product Limit Method (PLM), which has used in developing the probability of freeway breakdown, to predict the occurrence of an incident, as the mechanisms of breakdown and incidents have certain similarities. The objective of this chapter is to investigate whether incident occurrence (primarily crashes) can be detected based on likelihood of incident functions, and to compare the results to previously developed algorithms reported in the literature.

Likelihood of incident functions are developed using the PLM for seven different traffic parameters: average flow, standard deviation of flow, average occupancy, standard deviation of occupancy, speed difference (speed limit minus speed), standard deviation and average variation of speed. An index of incident detection is developed based on these functions. This chapter also compares the results to previously developed incident detection algorithms reported in the literature.

This chapter is structured as follows. The first section presents the data set used, and defines the operational parameters that will be used in predicting the incident potential as a function of traffic states. The second section presents the results of the PLM for both non-congested conditions and congested conditions, obtains the relationship calibrated between incident potential and each traffic parameter (linear or polynomial, etc.) by regression, and proposes the incident detection index. The third section evaluates the results based on detector data. The last section provides conclusions and recommendations.
Methodology

Incident occurrence is not deterministic, there is no unique combination of conditions that always lead to an incident. Traffic parameters also fluctuate significantly. Thus using probabilistic techniques (in this case the PLM) seems to be a promising approach in detecting incidents.

Based on literature review findings, several parameters have been considered in incident detection. These include occupancy, volume, speed, and standard deviation of occupancy. This chapter also considers several other parameters that have the potential to be correlated with incidents: 5-min standard deviation of flow, speed difference (speed limit minus speed), 5-min standard deviation of speed, and 5-min average variation of speed. A single parameter may not correlate completely with incident occurrence, however multiple parameters used at once may provide a better indication of an incident.

Likelihood of incident functions are developed using the PLM technique for each parameter, and an index of incident detection is developed based on these functions. The incident functions and the index are developed for non-congested and congested conditions separately, to investigate whether indices obtained by condition generate better results. This approach is different from previous research, wherein a single incident detection algorithm is developed for all traffic conditions or for different traffic states (e.g., McMaster). This decision is based on the observation that traffic operations under incident conditions have a similar pattern to that under recurrent congestion. It is expected that separate indices for different traffic conditions would decrease false detection of incidents during recurrent congestion.
This section illustrates the methodology developed to detect incidents, including the description of data used in the analysis, the definition of parameters that may correlate with the likelihood of an incident and the definition of traffic states, and the calculation of the PLM functions.

**Description of Data**

Data collected from three sites are used in this chapter: Toronto, Minneapolis, and Portland, as they have more incidents verified. All crashes with durations of 5 minutes or longer are used, and are classified into two groups: crashes occurring before congestion (congestion is caused by incidents), and crashes occurring during congested conditions (congestion already exists when incident occurs).

Table 7-1 provides an overview of the study site characteristics for the sites used in this chapter, and the respective data used in the analysis.

**Definition of Parameters and Traffic States**

Based on the discussion provided earlier, a total of seven parameters are used for incident detection: flow, 5-min standard deviation of flow, occupancy, 5-min standard deviation of occupancy, speed difference (speed limit minus speed), 5-min standard deviation of speed, and 5-min average variation of speed. Flow and occupancy are obtained every one minute.

The standardized speed difference was developed to take into consideration the differences in the speed limits between sites. It is calculated as,

\[ v_d = \frac{(V_{\text{limit}} - v_i)}{V_{\text{limit}}} \]  

(Eq. 7-1)

Where,

\[ V_{\text{limit}} \] is the speed limit at the data collection site (mph)
\( \nu_i \) is the speed at interval i (mph)

Traffic operation is divided into non-congested conditions and congested conditions based on speed, because usually a drop in speed is the first indication of congestion occurrence.

Based on the analysis of non-incident data at Toronto site, the average 1-minute speed at the beginning of recurrent congestion is about 43 mi/h. Thus, a critical speed of 43 mi/h is used in dividing traffic into non-congested and congested conditions. Operation with speed higher than 43 mi/h is considered to be non-congested conditions.

For the Minneapolis and Portland sites, the average 1-minute speeds at the beginning of recurrent congestion are about 40.7 mi/h and 37.4 mi/h respectively. Thus, a critical speed of 41 mi/h and 37 mi/h respectively are used in dividing traffic into non-congested and congested conditions at these two sites. The critical speed at Portland site is about 3 mph lower than that at Minneapolis site, probably because the speed limit at Portland site (55 mph) is lower than that at Minneapolis site (60 mph).

**Calculation of the PLM**

The likelihood of incident curves are developed using the PLM for each parameter at each site, for non-congested and congested conditions separately. In applying the PLM, the likelihood of an incident is calculated using equations 6-3 through 6-6, except that the number of breakdowns is replaced by the number of incidents. The incident interval is the one immediately preceding an incident.

At each site, crashes before congestion and those during congestion are used in developing the PLM models for non-congested and congested conditions respectively. All crashes with a duration of 5-min or longer except two crashes, which will be used in
evaluation later, are used in applying the PLM, based on each of the seven parameters illustrated above. The following steps are followed:

1) For each crash, obtain the average 1-minute parameter for 10 minutes prior to the crash. Use the 10th minute parameters as the parameters that resulted in incident.

2) Obtain the same amount of data points from days without crashes during similar time periods and at the same detector locations.

3) Calculate the likelihood of an incident using equations 6-3 through 6-6.

4) Develop the likelihood distribution curve as a function of the subject parameter.

**Results of the PLM and Incident Detection Index**

This section presents the results of the PLM models, which predict the potential of incident occurrence, and the proposed incident detection indices for both non-congested and congested conditions.

**PLM for Non-congested Conditions**

Crashes occurring before congestion are used in applying the PLM for non-congested conditions. The likelihood of incident functions are estimated based on each of the seven traffic parameters. The results are shown in Figure 7-1 A) – G). In each of the graphs the vertical axis indicates the likelihood of an incident in a one-minute interval. In these graphs, each data point represents one incident corresponding to the parameter values.

It is observed from Figure 7-1 that, the likelihood of incident has an exponential relationship with flow and occupancy, a polynomial relationship with the speed difference, and a linear relationship with the other parameters. As shown the trends are similar at the three sites, however there are some differences, especially for the parameters occupancy and speed difference. For example, the likelihood of incident based on occupancy is higher at the Portland site than at the other two sites. These
differences are due to reasons including different speed limits, different geometric characteristics (the number of on/off ramps downstream), and different number of lanes at the three sites. At each site, the mathematic relationship calibrated between likelihood of incident and each parameter is shown in Table 7-2.

**PLM for Congested Conditions**

Crashes occurring during congested conditions are used in applying the PLM for congested conditions. The likelihood of incident functions are estimated based on each of the seven traffic parameters. The results are shown in Figure 7-2 A) - G). In each of the graphs the vertical axis indicates the likelihood of an incident in a one-minute interval. In these graphs, each data point represents one incident corresponding to the parameter values.

It is observed from Figure 7-2 that, for congested conditions, the likelihood of incident has a logarithmic relationship with flow and occupancy, a polynomial relationship with the speed difference, and a linear relationship with the other parameters. These relationships are different from that under non-congested conditions, especially for flow and occupancy. Data observations show that, for congested conditions, the likelihood of incident decrease with occupancy and flow. As shown, there are also some differences in the relationships at the three sites. The reasons are similar to that for non-congested conditions. At each site, the calibrated mathematic relationship between likelihood of incident and each parameter is shown in Table 7-3.

**Proposed Index of Incident Detection**

To identify which of the seven parameters are significant in incident detection, six crashes (three occurring before congestion and three occurring during congestion) used in developing the PLM functions are selected at each site. For each crash, the likelihood
of an incident based on each of the seven parameters is plotted against time. Parameters positively or negatively correlated with incident occurrence during actual incident conditions are viewed to be significant in incident detection. Conversely, parameters that are positively or negatively correlated with non-incident conditions are also identified. Parameters that show no clear change between incident and non-incident conditions are not good indicators of incident occurrence and are thus not considered further.

Two examples of the plots from Toronto site are shown in Figure 7-3 and Figure 7-4. In Figure 7-3, the incident occurs before congestion. In Figure 7-4, the incident occurs during congestion.

As shown in Figure 7-3, the blue area in the middle indicates the incident duration 15:00-15:30, the two orange periods indicate the first 5-min of transition states (non-congested to congested states or congested to non-congested states). Three parameters flow, standardized speed difference, and 5-min average variation of speed, have peaks identical with the actual incident. A fourth parameter, average occupancy, has peaks contrary to the actual incident. The other three parameters show no clear change of incident likelihood between incident and non-incident conditions.

The similar patterns are found from Figure 7-4 and at the other two sites. Thus, the four parameters: flow, occupancy, standardized speed difference, and 5-min average variation of speed, are found to be significant in incident detection. The proposed index of incident detection is calculated as the sum of the likelihood of incident based on flow, standardized speed difference, and 5-min average variation of speed minus the likelihood based on occupancy.
The proposed index of incident detection for Toronto site is,

\[ INDEX = 1.5 \times P(\text{incident})_{flow} - 1.0 \times P(\text{incident})_{occ} + P(\text{incident})_{vd} + P(\text{incident})_{CVS} \]

The proposed index of incident detection for Minneapolis site is,

\[ INDEX = 2.0 \times P(\text{incident})_{flow} - 0.5 \times P(\text{incident})_{occ} + P(\text{incident})_{vd} + P(\text{incident})_{CVS} \]

The proposed index of incident detection for Portland site is,

\[ INDEX = 1.0 \times P(\text{incident})_{flow} - 1.0 \times P(\text{incident})_{occ} + P(\text{incident})_{vd} + P(\text{incident})_{CVS} \]

The coefficients before each of the parameters are obtained by trial and error. For example, when the coefficient of flow increases, the detection rate may increase and the false alarm rate may also increase. While when it decreases, the detection rate may decrease and the false alarm rate may decrease. Although the crashes used in developing the PLM curves are all with duration of greater than 5 minutes, a crash is detected to occur if the index exceeds 0.5 for a continuous duration of at least two minutes.

It is noted that the proposed incident detection indices for non-congested and congested conditions are in the same form. However, the PLM functions based on which they are developed are different.

**Model Evaluation and Comparison to Previous Research**

The proposed incident detection index is evaluated by detector data and compared to three previous models that are commonly used: California No. 7, California No. 8, and McMaster. A subset of data (including both crashes and non-incident days) collected from each of the three data collection sites are used to evaluate the proposed incident detection index. Information about the evaluation data at the three sites are summarized in the first four columns in Table 7-4.
The evaluation data are available in two methods: raw data, which is originally recorded each 20 seconds and aggregated every one minute, and clean data, which is also aggregated every one minute and processed by a program by removing missing or erroneous observations, etc., (additional information is provided in Elefteriadou et al., 2009). The proposed algorithm is used to predict the likelihood of incidents for non-congested conditions and congested conditions respectively, based on both clean data and raw data. As illustrated earlier, a crash is detected to occur if the index (total incident likelihood) is above 0.5 for a continuous duration of at least two minutes. The evaluation results are shown in Figure 7-5 A) - C). The detected incident time and duration are shown in the fifth column in Table 7-4.

It is observed from Figure 7-5 that all the crashes used in evaluation are detected 1-min or 2-min later than the actual crash starting time. The predicted crash durations are also mostly the same as the actual crash durations. However, there are some intervals the index (total incident likelihood) is above 0.5 with durations of 2 minutes or more, however, no incident is recorded to occur. These are viewed to be false alarms. The number of false alarms for each day for both clean data and raw data is summarized in the last column in Table 7-4. As shown, the clean data generates fewer false alarms than the raw data. It is also observed from Figure 7-5 that most of the false alarms occur during the AM periods, probably because there are more recurrent congestion during the AM periods.

The proposed incident detection index is compared to three previous models: California No. 7, California No. 8, and McMaster algorithm, based on the three measure
of effectiveness: DR, FAR, MTTD. The calculation of DR and MTTD are the same as previous research. As all the incidents are detected, the DR is 100%.

As shown in Table 7-4, the time differences between the incidents are detected and the incidents actually occur are: 1 min and 2 min at the Toronto site, 1 min and 1 min at the Minneapolis site, and both 2 min at Portland site. Thus MTTD is calculated as,

\[
MTTD_{\text{Toronto}} = (1 + 2)/2 = 1.5 \text{ (min)}
\]

\[
MTTD_{\text{Minneapolis}} = (1 + 1)/2 = 1.0 \text{ (min)}
\]

\[
MTTD_{\text{Portland}} = (2 + 2)/2 = 2.0 \text{ (min)}
\]

\[
MTTD = (1 + 2 + 1 + 1 + 2 + 2)/6 = 1.5 \text{ (min)}
\]

FAR is calculated as the ratio of the number of false alarms to the total number of algorithm applications (executed every minute in this chapter), which is also similar to previous research. The FAR for clean data is calculated as,

\[
FAR_{\text{Toronto}} = \frac{(3 + 0 + 0 + 0)/4 \text{ False Alarms}}{1 \text{ iteration/Minute} \times 60 \text{ Minutes/Hour} \times 24 \text{ hours}} \times 100\% = 0.05\%
\]

\[
FAR_{\text{Minneapolis}} = \frac{(2 + 3 + 1 + 2)/4 \text{ False Alarms}}{1 \text{ iteration/Minute} \times 60 \text{ Minutes/Hour} \times 19 \text{ hours}} \times 100\% = 0.18\%
\]

\[
FAR_{\text{Portland}} = \frac{(2 + 2 + 0 + 2)/4 \text{ False Alarms}}{1 \text{ iteration/Minute} \times 60 \text{ Minutes/Hour} \times 19 \text{ hours}} \times 100\% = 0.13\%
\]

\[
FAR = (0.05\% + 0.18\% + 0.13\%)/3 = 0.12\%
\]

The evaluation results of the proposed index (based on both clean data and raw data) are summarized in Table 7-5. As shown in Table 7-5, when compared to previous algorithms, the proposed index yields higher detection rate and lower mean detect time. The false alarm rates for clean data at Toronto and Portland site are lower than the California algorithms but a litter higher than the McMaster algorithms. The false alarm rates for raw data are higher than the three comparing algorithms. However, the average FAR at the three sites is within the acceptable levels desired by operators and quite low compared to other algorithms as shown in Table 2-7. The index only needs
operation data at one point, and thus is a preferable option in freeway incident detection application. It's also observed from Table 7-5 that the raw data generates the same detection rate and detection time as the clean data, but has higher FAR than the clean data. The average FAR based on the raw data at the three sites is still within the acceptable levels desired by operators.

As indicated earlier, most of the FAR occurs during the AM periods. When applying the indices separately for AM and PM periods, the results are quite different. For the AM periods, the FAR are 0.1%, 0.42%, and 0.26% at the Toronto, Minneapolis and Portland site respectively, and the average MTTD is 1.4 minutes at the three sites; For the PM periods, the FAR are 0.0%, 0.0%, and 0.04% at the Toronto, Minneapolis and Portland site respectively, and the average MTTD is 2.0 minutes. The values of DR are all 100%. This suggests that the proposed indices work much better during the PM than during the AM.

Alternately, the FAR can be reduced by decreasing the coefficients before each of the parameters in the proposed incident detection index. For example, for Minneapolis site with clean data, when the coefficient before flow is decreased from 2 to 1.5 and the coefficient before occupancy is decreased from -0.5 to -1.0, the index becomes (alternate index),

\[ INDEX = 1.5 \times P(\text{incident})_{flow} - 1.0 \times P(\text{incident})_{occ} + P(\text{incident})_{vd} + P(\text{incident})_{CVS} \]

The resulted FAR is 0.13%, lower than the original value of 0.18%. However, one incident cannot be detected, making the DR at the Minneapolis site as 50% and the average DR at the three sites as 83.3% (= 5/6). Thus, we recommend that it is better to have higher detection rate than a very low false alarm rate.
Concluding Remarks

The PLM has been used in previous research to estimate the probability of breakdown. This chapter uses this method to estimate the likelihood of incident (particularly crashes) and then detect incident. Data are collected from three freeways in North America. The PLM functions were first developed at each site to estimate the incident potential for both non-congested and congested conditions, based on seven traffic parameters. An incident detection index is then proposed based on the PLM functions, considering parameters that correlate well with the likelihood of an incident.

The following conclusions are drawn:

• For both non-congested and congested conditions, four factors are found to be significant in incident detection. Flow, standardized speed difference, and 5-min average variation of speed are positively correlated with incidents, as the likelihood of incident based on these parameters increases when an incident occurs. A fourth parameter, occupancy, is negatively correlated with the likelihood of an incident. These four parameters are included in the incident detection index.

• The proposed incident detection index are evaluated by detector data collected from each site and compared to several previous algorithms. The results show that the proposed index generates higher detection rate and lower mean detect time.

• The clean data generates fewer false alarms than the raw data.

• Most of the false alarms occur during the AM periods, probably because there are more recurrent congestion during the AM periods. When applying the indices separately for AM and PM periods, the results suggest that the proposed indices work much better during the PM than during the AM.

• The false alarm rate can be reduced by decreasing the coefficients before each of the parameters in the proposed incident detection index. However, the detection rate might also be reduced.

The incident detection index proposed in this dissertation is based on data from one detector at a time. Such an index could be very helpful in traffic management of freeway facilities. The PLM models should be developed based on data from the subject sensor, as each site generates a slightly different set of curves. Further research is
needed to determine the relationship between these curves and the prevailing traffic and environmental conditions at each site. Additional validation of the proposed model would be useful to confirm that the set of indices proposed here can be applied in a wide variety of situations.
Table 7-1. Overview of site characteristics and data

<table>
<thead>
<tr>
<th>Site</th>
<th>Location</th>
<th>Section length (mi)</th>
<th>Total # of lanes</th>
<th>Speed limit (mph)</th>
<th>Data period</th>
<th># of non-incident days</th>
<th># of crashes before congestion</th>
<th># of crashes during congestion</th>
</tr>
</thead>
<tbody>
<tr>
<td>QEW</td>
<td>Toronto, Canada</td>
<td>6.5</td>
<td>3</td>
<td>60</td>
<td>01/2005-12/2005</td>
<td>55</td>
<td>18 (1)</td>
<td>9 (1)</td>
</tr>
<tr>
<td>I-494</td>
<td>Minneapolis, MN</td>
<td>3</td>
<td>2</td>
<td>60</td>
<td>09/2006-08/2007</td>
<td>33</td>
<td>17 (1)</td>
<td>11 (1)</td>
</tr>
<tr>
<td>OR</td>
<td>Portland, OR</td>
<td>7</td>
<td>2</td>
<td>55</td>
<td>11-12/2007</td>
<td>32</td>
<td>33 (1)</td>
<td>14 (1)</td>
</tr>
</tbody>
</table>

Note: numbers in the parentheses indicate the number of data points used for evaluation.

Table 7-2. Relationship between likelihood of incident and each parameter for non-congested conditions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Toronto</th>
<th>Minneapolis</th>
<th>Portland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow</td>
<td>0.0009<em>exp(0.0022</em>flow)</td>
<td>0.0004<em>exp(0.0028</em>flow)</td>
<td>0.0006<em>exp(0.0029</em>flow)</td>
</tr>
<tr>
<td>5-min std.flow</td>
<td>0.0005*std.flow-0.0636</td>
<td>0.0005*std.flow-0.0511</td>
<td>0.0005*std.flow-0.0426</td>
</tr>
<tr>
<td>Occ</td>
<td>0.0011<em>exp(0.3011</em>occ)</td>
<td>0.0007<em>exp(0.2803</em>occ)</td>
<td>0.0026<em>exp(0.2681</em>occ)</td>
</tr>
<tr>
<td>5-min std.occ</td>
<td>0.0525*std.occ-0.0576</td>
<td>0.044*std.occ-0.0447</td>
<td>0.0656*std.occ-0.0544</td>
</tr>
<tr>
<td>Speed</td>
<td>1.0039<em>vd^2+0.1766</em>vd</td>
<td>0.9672<em>vd^2+0.4711</em>vd</td>
<td>0.8019<em>vd^2+0.1742</em>vd</td>
</tr>
<tr>
<td>difference</td>
<td>+0.0313</td>
<td>+0.0691</td>
<td>+0.0274</td>
</tr>
<tr>
<td>5-min std.v</td>
<td>0.0241*std.v-0.0348</td>
<td>0.0257*std.v-0.0898</td>
<td>0.0247*std.v-0.0338</td>
</tr>
<tr>
<td>5-min cvs</td>
<td>1.0327*cvs-0.0046</td>
<td>1.5681*cvs-0.0762</td>
<td>1.0105*cvs-0.0178</td>
</tr>
</tbody>
</table>

Table 7-3. Relationship between likelihood of incident and each parameter for congested conditions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Toronto</th>
<th>Minneapolis</th>
<th>Portland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow</td>
<td>2.0453-0.27*LN(flow)</td>
<td>2.911-0.386*LN(flow)</td>
<td>2.9509-0.392*LN(flow)</td>
</tr>
<tr>
<td>5-min std.flow</td>
<td>0.0004*std.flow-0.0435</td>
<td>0.0007*std.flow-0.0711</td>
<td>0.001*std.flow-0.179</td>
</tr>
<tr>
<td>Occ</td>
<td>0.8367-0.245*LN(occ)</td>
<td>0.8495-0.244*LN(occ)</td>
<td>0.9783-0.31*LN(occ)</td>
</tr>
<tr>
<td>5-min std.occ</td>
<td>0.0113*std.occ-0.0109</td>
<td>0.0171*std.occ-0.0203</td>
<td>0.0147*std.occ-0.0239</td>
</tr>
<tr>
<td>Speed</td>
<td>0.4491*vd^2-</td>
<td>0.8296*vd^2-</td>
<td>0.7691*vd^2-</td>
</tr>
<tr>
<td>difference</td>
<td>0.1478*vd+0.0036</td>
<td>0.2905*vd+0.0286</td>
<td>0.3466*vd+0.0466</td>
</tr>
<tr>
<td>5-min std.v</td>
<td>0.0161*std.v-0.0451</td>
<td>0.0131*std.v-0.0324</td>
<td>0.0263*std.v-0.1053</td>
</tr>
<tr>
<td>5-min cvs</td>
<td>0.3664*cvs-0.029</td>
<td>0.9095*cvs-0.0944</td>
<td>0.8045*cvs-0.1073</td>
</tr>
</tbody>
</table>
### Table 7-4. Characteristics of evaluation data and evaluation results

<table>
<thead>
<tr>
<th>Site</th>
<th>Date</th>
<th>Incident time</th>
<th>Incident type*</th>
<th>Detected incident time</th>
<th># False alarms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto</td>
<td>Jul 25 2005</td>
<td>7:14-7:39</td>
<td>During congestion</td>
<td>7:15-7:36</td>
<td>3 (clean), 9 (raw )</td>
</tr>
<tr>
<td></td>
<td>Jul 28 2005</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0 (clean), 0 (raw )</td>
</tr>
<tr>
<td></td>
<td>Aug 15 2005</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0 (clean), 0 (raw )</td>
</tr>
<tr>
<td></td>
<td>Aug 17 2005</td>
<td>18:12-19:09</td>
<td>Before congestion</td>
<td>18:14-19:09</td>
<td>0 (clean), 0 (raw )</td>
</tr>
<tr>
<td></td>
<td>Sep 11 2006</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>2 (clean), 3 (raw )</td>
</tr>
<tr>
<td></td>
<td>Sep 12 2006</td>
<td>7:56-8:51</td>
<td>During congestion</td>
<td>7:57-8:00 &amp; 8:26-8:45</td>
<td>3 (clean), 5 (raw )</td>
</tr>
<tr>
<td></td>
<td>Sep 14 2006</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>1 (clean), 1 (raw )</td>
</tr>
<tr>
<td></td>
<td>Nov 30 2007</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>2 (clean), 4 (raw )</td>
</tr>
<tr>
<td></td>
<td>Dec 11 2007</td>
<td>17:13-18:07</td>
<td>Before congestion</td>
<td>17:15-18:07</td>
<td>0 (clean), 0 (raw )</td>
</tr>
<tr>
<td></td>
<td>Dec 12 2007</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>2 (clean), 3 (raw )</td>
</tr>
</tbody>
</table>

* Whether incident occurs before congestion or during congested conditions.

### Table 7-5. Comparison results of the proposed index to previous algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Site</th>
<th>Detection Rate</th>
<th>False Alarm Rate</th>
<th>Mean Time to Detect (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>California No. 7</td>
<td></td>
<td>67%</td>
<td>0.134%</td>
<td>2.91</td>
</tr>
<tr>
<td>California No. 8</td>
<td></td>
<td>68%</td>
<td>0.177%</td>
<td>3.04</td>
</tr>
<tr>
<td>McMaster</td>
<td></td>
<td>68%</td>
<td>0.0018%</td>
<td>2.2</td>
</tr>
<tr>
<td>Proposed index</td>
<td>Toronto</td>
<td>100%</td>
<td>0.05%</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Minneapolis</td>
<td>100%</td>
<td>0.18%</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Portland</td>
<td>100%</td>
<td>0.13%</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>100%</td>
<td>0.12%</td>
<td>1.5</td>
</tr>
<tr>
<td>Proposed index</td>
<td>Toronto</td>
<td>100%</td>
<td>0.16%</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Minneapolis</td>
<td>100%</td>
<td>0.24%</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Portland</td>
<td>100%</td>
<td>0.29%</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>100%</td>
<td>0.23%</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Figure 7-1. Likelihood of incident for non-congested conditions. A) Flow-based model, B) 5-min std. flow-based model, C) Occupancy-based model, D) 5-min std. occ-based model, E) Speed difference-based model, F) 5-min std.v-based model and G) 5-min cvs-based model.
Figure 7-2. Likelihood of incident for congested conditions. A) Flow-based model, B) 5-min std. flow-based model, C) Occupancy-based model, D) 5-min std. occ-based model, E) Speed difference-based model, F) 5-min std. v-based model and G) 5-min cvs-based model.
Figure 7-3. Likelihood of incident potential based on each parameter for incidents occurring before congestion
Figure 7-4. Likelihood of incident potential based on each parameter for incidents occurring during congestion.
Figure 7-5. Evaluation results of incident detection. A) at Toronto site, B) at Minneapolis site and C) at Portland site.
CHAPTER 8
FUTURE WORK

This dissertation conducts analysis on the impact of incidents on freeway flow from three aspects: impact on operational conditions, impact on capacity, and impact on congestion, and detects incidents from operational conditions. In the dissertation, there are some limitations on the dataset and analysis method used. To make the findings more general and results more practical, the following areas are proposed for further research:

(1) This dissertation considered capacity for shoulder plus one lane or two lanes affected conditions, however the number of data points was not adequate for those to be included in the model development. Thus it is recommended that additional data be collected for those types of incidents.

(2) With respect to incident data collection, ideally incidents should be reported such that the number of lanes closed throughout the incident is reported. The Portland, Oregon database provides such an incident log, which greatly facilitates the modeling of incident capacity.

(3) The number of lanes closed by incidents is important for incident capacity analysis. For sites without information about the number of lanes affected/closed by incidents, it is suggested to predict the probability of lane closure from incident type and the number of vehicles involved in incidents. The impacts of geographic characteristics on operational conditions can be obtained by comparing the results at different sites.

(4) It would be better to estimate the probability of incident-induced breakdown and detect incidents based on the combination of several parameters. It is recommended to identify regions with high probability of incident and incident-induced breakdown.

(5) Further research is necessary to detect the exact location of the incident. This could be achieved by using occupancy data from both upstream and downstream stations. The location of an incident would have increasing occupancy upstream and decreasing occupancy downstream.

(6) Develop guidelines for freeway management based on the results obtained in this dissertation. Further research on the development of ramp management strategies responsive to incidents is necessary. It is recommended to consider the capacity reduction caused by incidents, breakdown probability and incident detection in ramp metering strategy. For example, for traffic regions with high probabilities of
incident and incident-induced breakdown, the metering rate should be set at a lower value. It is also recommended to taken into consideration the location of incidents in ramp metering. For instance, when incident occurs downstream of the bottleneck, the metering rate upstream should decrease, while when incident occurs upstream of the bottleneck, the metering rate might increase.

(7) The finding also can be useful in implementing Variable Speed Limit algorithms to optimize freeway operations.
APPENDIX A
FORMAT OF DATA

Appendix A describes the format of data at each site.

• Interstate 15 SB

Traffic data at this site are put in comma separated text files with the format of d11_text_station_raw_ YYYY_MM_DD.txt. Each file contains information for all the detectors per day. Thus it needs to extract data for each detector from the file. The time interval is 30 seconds. The contents of the data fields are shown in Table A-1.

• Interstate 5 NB

Traffic data at this site are put in comma separated text files with the format of d3_text_station_raw_ YYYY_MM_DD.txt. Each file contains information for all the detectors per day. Thus it needs to extract data for each detector from the file. The time interval is 30 seconds. The contents of the data fields are shown in Table A-2.

• Queen Elizabeth Way (QEW, Toronto, Canada)

Traffic data at this site includes mainline detector data and ramp detector data, which are comma separated text files with the format of YYYY-MM-DD.txt. The time interval is 20 seconds. The contents of the data fields are shown in Table A-3.

There are totally 4,322 lines in each data file, as the data are collected 24 hours per day.

• US 217 SB, Portland, Oregon

Volume, occupancy, and speed data at this site are saved in separated files, each file carries information for one detector per day. The available information concludes time and occupancy or volume or speed. The time interval is 20 seconds. The format of the data is shown in Table A-4.
● US 494 EB, Minnesota

Traffic data at this site are saved separated file, each file carries information for all the detectors per day. The available information concludes time, occupancy, speed and volume. The time interval is 1 minute. The format of the data is shown in A-5.
Table A-1. Format of data at I-15 SB

<table>
<thead>
<tr>
<th>Col. #</th>
<th>Column / Field</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Timestamp</td>
<td>MM/DD/YYYY HH24:MI:SS,</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Detector number</td>
<td></td>
<td>1108509,</td>
</tr>
<tr>
<td>3</td>
<td>Lane 1 Flow</td>
<td>Veh/30 sec</td>
<td>14,</td>
</tr>
<tr>
<td>4</td>
<td>Lane 1 Occupancy</td>
<td>%</td>
<td>.0978,</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Lane 2 Flow</td>
<td>Veh/30 sec</td>
<td>11,</td>
</tr>
<tr>
<td>7</td>
<td>Lane 2 Occupancy</td>
<td>%</td>
<td>.0833,</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Lane 8 Flow</td>
<td>Veh/30 sec</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Lane 8 Occupancy</td>
<td>%</td>
<td></td>
</tr>
</tbody>
</table>

Table A-2. Format of data at I-5 NB

<table>
<thead>
<tr>
<th>Col. #</th>
<th>Column / Field</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Timestamp</td>
<td>MM/DD/YYYY HH24:MI:SS,</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Detector number</td>
<td>MM/DD/YYYY HH24:MI:SS,</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Lane 1 Flow</td>
<td>Veh/30 sec</td>
<td>314886</td>
</tr>
<tr>
<td>4</td>
<td>Lane 1 Occupancy</td>
<td>%</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>.0978,</td>
</tr>
<tr>
<td>6</td>
<td>Lane 2 Flow</td>
<td>Veh/30 sec</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Lane 2 Occupancy</td>
<td>%</td>
<td>11,</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Lane 8 Flow</td>
<td>Veh/30 sec</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Lane 8 Occupancy</td>
<td>%</td>
<td></td>
</tr>
</tbody>
</table>

Table A-3. Format of data at the Queen Elizabeth Way (QEW) site

<table>
<thead>
<tr>
<th>Col. #</th>
<th>Field Name</th>
<th>Units</th>
<th>Example field entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Timestamp</td>
<td></td>
<td>2005-05-16 18:00:20</td>
</tr>
<tr>
<td>2</td>
<td>Total Lane Volume</td>
<td>Veh/h</td>
<td>6840.018</td>
</tr>
<tr>
<td>3</td>
<td>Ave. Lane Occupancy</td>
<td>%</td>
<td>14.3333</td>
</tr>
<tr>
<td>4</td>
<td>Ave. Lane Speed</td>
<td>Km/h</td>
<td>90.6667</td>
</tr>
</tbody>
</table>
### Table A-4. Format of data at the US 217 SB

<table>
<thead>
<tr>
<th>Start time</th>
<th>Occupancy/Volume/Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006-04-03 00:00:00-07</td>
<td>0.5 / 90 / 51</td>
</tr>
<tr>
<td>2006-04-03 00:00:20-07</td>
<td>0 / 0 / 0</td>
</tr>
<tr>
<td>2006-04-03 00:00:40-07</td>
<td>3.5 / 360 / 42</td>
</tr>
<tr>
<td>2006-04-03 00:01:00-07</td>
<td>0.5 / 90 / 40</td>
</tr>
</tbody>
</table>

### Table A-5. Format of data at the US 494 EB

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12:01 AM</td>
<td>3.472222</td>
<td>29.45455</td>
<td>3</td>
<td>.</td>
<td>3.444444</td>
<td>21.7742</td>
</tr>
<tr>
<td>12:02 AM</td>
<td>0.361111</td>
<td>94.4056</td>
<td>1</td>
<td>.</td>
<td>0.555556</td>
<td>67.49999</td>
</tr>
<tr>
<td>12:03 AM</td>
<td>0.333333</td>
<td>102.2727</td>
<td>1</td>
<td>.</td>
<td>0.583333</td>
<td>64.28571</td>
</tr>
</tbody>
</table>
LIST OF REFERENCES


Balke, K.N. An evaluation of existing incident detection algorithms. Research Report, FHWA/TX-93/1232-20, Texas Transportation Institute, the Texas A&M University System, College Station, TX, November 1993.


Kloeden CN, McLean AJ, Moore VM, Ponte G. Traveling Speed and the Risk of Crash Involvement. NHMRC Road Accident Research Unit. The University of Adelaide - November 1997.


Liu P. Freeway incident likelihood prediction models: development and application to traffic management systems. Purdue University, 1997.


BIOGRAPHICAL SKETCH

Cuie Lu was born in Anlu, Hubei, China. She joined in Tongji University in 2000, received her Bachelor of Engineering in traffic engineering in 2004 and Master of Engineering in traffic information engineering & control in 2007. She was a research assistant in transportation engineering since August 2007 in University of Florida. Her research topics include traffic safety management, operational analysis, traffic modeling.