

ABSTRACT

Title of Dissertation: DEVELOPMENT OF A TRAFFIC INCIDENT
MANAGEMENT SUPPORT SYSTEM

Minsu Won, Doctor of Philosophy, 2019

Dissertation directed by: Gang-Len Chang, Professor,
Department of Civil & Environmental Engineering

Highway incidents, primary contributors to traffic congestion, often cause not only significant delays to the daily roadway users, but also the reliability of transportation systems and even the productivity of the supply-chains of some vital industry sectors. To mitigate the impacts of such incidents and to recover the performance of highway systems as safely and quickly as possible, it is essential that the responsible highway agencies shall operate an efficient system to detect an incident, estimate the required clearance duration, assess the resulting traffic impacts, and then take necessary control actions.

To do so, the most critical task is to have a reliable estimate of a detected incident's impacts. However, providing the information of time-varying incident impacts to the general public at the desirable level of accuracy is a challenging task due to difficulties in having sufficient data and the complex relations between key factors contributing to the incident impacts. The purpose of this study is to develop a traffic incident management (TIM) support system which is capable of providing robust and reliable information with

respect to the estimated clearance duration and capacity drop of a detected incident and its temporal as well as spatial evolution of traffic impact patterns.

The proposed incident management support system consists of two main components, one for estimating the incident duration and the other for computing the resulting capacity drop of the roadway segment plagued by the incident. The first component serves to provide a robust incident duration estimate, using several specially-designed methods to effectively tackle the unique distribution patterns of the incident duration data and their complex correlations among contributing factors, which includes classification model, continuous model, and supplemental rules to first produce an initial interval estimate, and then a point estimate along with the outlier information. The second component is designed to estimate the additional roadway capacity reduction (i.e., capacity drop) due to the lane blockages by incidents and the response operations, allowing the control center operators to assess the spatial and temporal incident impacts on the highway network, and take necessary control actions in a timely manner.

The proposed TIM support system has the following key features: 1) providing the initial estimate of incident duration, based on limited data available at the early stage of incident responses and operations; 2) updating the estimated incident duration with a specially-design process and models when more data become available; 3) implementing an integrated estimation methodology to circumvent the variances due to the unique characteristics associated with recorded incident data (e.g., highly skewed distribution, complex correlations among the explanatory variables, mixed qualitative and quantitative variables, and heteroscedasticity); 4) generating the estimated additional capacity reduction for the highway segment plagued by the lane-closure and lane-changing activities during

the incident clearance operations using a reliable and trackable analytical model; and 5) providing a convenient and effective computation process to estimate time-varying incident impacts, such as queues and delays, in the highway network for real-time applications.

DEVELOPMENT OF A TRAFFIC INCIDENT MANAGEMENT SUPPORT SYSTEM

by

Minsu Won

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2019

Advisory Committee:

Professor Gang-Len Chang, Chair

Professor Martin Dresner, Dean's Representative

Professor Ali Haghani

Associate Professor Qingbin Cui

Associate Professor Shanjiang Zhu at the George Mason University

© Copyright by
Minsu Won
2019

Dedication

To my wife, Minyoung, and my son, Minjun,
for their love and support.

Acknowledgements

First of foremost, I would like to express my deepest sense of gratitude and indebtedness to my advisor, Dr. Gang-Len Chang, for his enlightening guidance and persistent inspiration throughout my Ph.D. study at the University of Maryland at College Park. My career has been strongly influenced by his dedication to excellent research work, and I have received a great deal of support and assistance throughout the writing of this dissertation.

I would like to acknowledge the members of my doctoral examination committee, Dr. Martin Dresner, Dr. Ali Haghani, Dr. Qingbin Cui, Dr. Cinzia Cirillo, and Dr. Shanjiang Zhu, Special Member at the George Mason University for their valuable comments and suggestions on my research work. My thanks also go to the faculty members and staff in our department, Dr. Paul M. Schonfeld, Dr. Lei Zhang, and Assistant Director Heather Stewart; and my colleagues in Traffic Safety and Operations Lab, Hyeonmi Kim, Yao Cheng, Yen-Hsiang Chen, and Yaou Zhang for their technical and moral supports in proceeding my research and projects.

I profoundly thank my beloved wife and son, Minyoung You and Minjun Won, who always believe in and provide me with a constant source of love, comfort, and encouragement with unconditional love. Their patience and understanding during my research are crucial for me to finish my study successfully. Finally, to my all family members, you deserve my deepest gratitude for your dedication and support during my Ph.D. study.

Table of Contents

Dedication	ii
Acknowledgements	iii
Table of Contents.....	iv
List of Tables	vi
List of Figures	vii
Chapter 1. Introduction	1
1.1. Research background	1
1.2. Objectives and scope of research	2
1.3. Organization of the dissertation proposal.....	3
Chapter 2. Literature review	7
2.1. Estimation of incident duration	7
2.2. Estimation of capacity reduction.....	23
Chapter 3. Structure of the traffic incident management (TIM) support system.....	38
3.1. Introduction	38
3.2. Key components of a TIM system	39
3.3. Key models to support a TIM system’s operations.....	42
Chapter 4. Interval-based model for incident duration	48
4.1. Introduction	48
4.2. Data description and preprocessing.....	49
4.3. Modeling concept and methodology	57
4.4. Model development and estimation results	58
4.5. Model evaluation and discussion	66
Chapter 5. Point-based model for incident duration	69
5.1. Modeling concept and methodology	69
5.2. Model development.....	69
5.3. Estimation results	77
5.4. Model evaluation and discussion	82
Chapter 6. Supplemental models for outliers and integration.....	85
6.1. Identifying outliers among detected incidents.....	85
6.2. An integrated estimation process	102
Chapter 7. Capacity analysis for a lane-closure highway segment: single-lane-open scenarios	109
7.1. Introduction	109
7.2. Model development.....	112
7.3. Numerical analysis of the model properties	125
7.4. Interim findings	129
Chapter 8. A generalized model for a lane-closure highway segment: multi-lane-open scenarios	130
8.1. Model development.....	130
8.2. Numerical analysis of the model properties	137

8.3. Experimental Validation.....	141
8.4. Interim findings	148
Chapter 9. Assessment of incident impacts and conclusions.....	150
9.1. Introduction	150
9.2. Assessment of incident impacts	151
9.3. Example of system applications	156
9.4. Conclusions and future research.....	165
Bibliography	170

List of Tables

Table 1. List of key factors associated with the response/clearance of a detected incident.....	50
Table 2. Sample sizes by lane blockage status and years (Collision).....	51
Table 3. Accuracy of interval estimates by incident types and # of blocked travel lanes	67
Table 4. Accuracy of interval estimates by CT categories	67
Table 5. Distribution of interval estimation errors by CT categories.....	67
Table 6. Summary of Classification Results with the Proposed Methodology.....	78
Table 7. Model Estimation Results for Each Group	79
Table 8. Estimated Weights for Each Group Produced from Robust Regression.....	81
Table 9. Accuracy (MAE) of point estimates by CT categories	83
Table 10. Precision (SD of error) of point estimates by CT categories	83
Table 11. Estimation Results of the PCA.....	94
Table 12. Quantile values for the ensemble outlier scores.....	96
Table 13. Accuracy (percentage) of interval estimates by CT categories	98
Table 14. Accuracy (MAE) of point estimates by CT categories	99
Table 15. Precision (SD of error) of point estimates by CT categories	99
Table 16. Improvement of point and interval estimates after the rules of outliers .	100
Table 17. Accuracy (percentage) of interval estimates by CT categories	104
Table 18. Accuracy (MAE) of point estimates by CT categories	104
Table 19. Precision (SD of error) of point estimates by CT categories	105
Table 20. Acceptable accuracy of point estimates by CT categories	105
Table 21. Distribution of point estimation errors by CT categories (Training set).	106
Table 22. Distribution of point estimation errors by CT categories (Test set).....	106
Table 23. Notations for a capacity drop model	113
Table 24. Variables and parameters for the numerical analyses	125
Table 25. Variables and parameters for the numerical analyses	137
Table 26. Effective capacities estimated by the proposed model under different conditions.....	137
Table 27. List of incidents.....	142
Table 28. Outflows from the lane closure by vehicle type	142
Table 29. Average merging times by vehicle type	142
Table 30. Estimated traffic conditions of each site using the nearest detectors.....	144
Table 31. Comparison of the observations and model estimates	145
Table 32. Notations for time-varying traffic queue estimation process	153
Table 33. Basic input information for the application example.....	161
Table 34. Incoming traffic flow rate from the detectors near the scene (Sep. 2, 2017)	161
Table 35. An example of output produced from the proposed incident management system	162

List of Figures

Figure 1. Dissertation proposal organization	6
Figure 2. Components of traffic incident duration [4].....	8
Figure 3. Distribution of the incident clearance time for MD I-95 in 2012-2017 (Collision only) [10]	9
Figure 4. Example of the capacity drop in the fundamental diagram [97]	23
Figure 5. Timeline for the sequence of operations in the TIM process [112].....	39
Figure 6. Framework of the proposed TIM support system	42
Figure 7. Components of the proposed incident duration prediction system	44
Figure 8. Framework of the proposed incident duration prediction system	45
Figure 9. Framework of the proposed model for estimating the capacity reduction due to traffic incidents	46
Figure 10. Study scope.....	49
Figure 11. Incident clearance time and frequency by year, incident type, hour, weekend, holiday, season, and pavement condition (Collison with travel lane blockage).....	51
Figure 12. Incident clearance time and frequency by county, operation center, lane blockage, involved vehicle, vehicle state, and response units (Collision with travel lane blockage)	52
Figure 13. (a) Distribution of time differences between the all-lane-reopen times and the event-cleared times, (b) distribution of time differences between the travel- lane-reopen times and the all-lane-reopen times.	56
Figure 14. Distribution of clearance times after the data preprocessing.....	56
Figure 15. Initial incident categorization and estimated clearance duration	59
Figure 16. An example of the classification rules for “< 30 mins” and “≥ 30 mins” in the CPI with two blocked travel lanes	62
Figure 17. An example of application for CPI with two-travel-lane blockage.....	64
Figure 18. Classification rules and incident clearance time ranges in the CF with travel lane blockage	66
Figure 19. Graphically illustration of the modeling concept and its procedures	71
Figure 20. Distribution and Q-Q plot (e.g., subgroup for “TOW arrived” cases)	75
Figure 21. Rule-based models for incidents having a small sample size (i.e., fatality, hazard material, and vehicle jack-knifed cases)	80
Figure 22. Flowchart of the modeling concept	87
Figure 23. Silhouette values to determine the optimal number of clusters for PAM 95	
Figure 24. Distributions of the outlier scores for each outlier detection algorithm..	96
Figure 25. Outlier score information to determine the threshold of outliers	97
Figure 26. Classification rules for the outliers.....	98
Figure 27. Examples when a point estimate and an interval estimate are mismatched	102
Figure 28. Decision tree model to adjust mismatched results (test set).....	103
Figure 29. Example of the capacity drop in one-out-of-two-lanes-blockage scenario	109

Figure 30. Illustration of a single-lane-open scenario and the key factors	112
Figure 31. Illustration of a relation between the initial void and effective void.....	116
Figure 32. Framework for developing a capacity-drop model	117
Figure 33. Illustration of effective merging boundaries	121
Figure 34. Illustration of interaction between an existing void and a vehicle merging ahead of its location	122
Figure 35. Effective voids vs. merging location with and without interactions between a void and vehicles merging ahead of its location	125
Figure 36. Sensitivity analysis of four variables on the capacity drop	127
Figure 37. Illustration of merging behaviors in the multi-lane-open scenario.....	130
Figure 38. Illustration of a double-lane occupied process on two open lane by one merging vehicle.....	132
Figure 39. Discretionary and mandatory lane changes and key variables	133
Figure 40. Key variables to estimate the effective capacity of the rightmost lane .	134
Figure 41. Key variables to estimate the effective capacity of the middle lane	135
Figure 42. Sensitivity analysis of different parameters on the capacity drop	139
Figure 43. Examples for RITIS for incident managements	141
Figure 44. Scatter plots and regression results to estimate a fundamental diagram for each site.....	144
Figure 45. Effective capacity with different merging ratio and SD of acceleration rate in each incident scene	146
Figure 46. Queuing diagram through three phases	152
Figure 47. Example of TIM support system application	157
Figure 48. Time-varying queue distance up to each incident phase	164
Figure 49. Example of the queue propagation information on the map	164

Chapter 1. Introduction

1.1. Research background

Over the past two decades, most metropolitan areas in the U.S. have been plagued by traffic congestion, especially during the peak commuting period. For example, in 2014, it has been reported that congestion caused urban commuters to travel extra 6.9 billion hours, 3.1 billion more gallons of fuel, and a total cost of \$160 billion [1]. Roughly half of the congestion delays experienced by U.S. travelers are non-recurring in nature, caused mainly by incidents or work zones that result in temporary traffic disruptions and take away part of the roadway capacity. The top three most common types of contributors to non-recurrent congestion, as reported by the Federal Highway Administration (FHWA), are incidents, including a flat tire, an overturned truck carrying hazardous material (25% of congestion), work-zone activities (10% of congestion), and inclement weather (15% of congestion) [2].

To contend with such non-recurring congestion, caused mainly by various types of incidents in daily commuting environments, most responsible highway agencies over the past decades have implemented various response plans and coordinated management systems, ranging from providing the information of incident location to the motorists via the Advanced Traveler Information System (ATIS) to deploying full-scale detouring operations with proper control strategies. Such incident management systems, depending on their embedded technical components, can provide the estimated maximum impact area, evolution of the traffic queues, predicted travel times over different roadway segments, resulting delays for en-route motorists, and suggested detour strategies [3]. However, providing the essential information to the roadway users at the acceptable reliability is a quite challenging task due to the difficulties associated with the data available and collected

by the traffic control center in the real-time response process, and the complex relations between those factors contributing to the incident impacts. Also, the demand for having high-quality information in a timely manner has increased significantly in the recent decade due to the advance in various communication technologies. Most en-route drivers would like to have more information when encountering incidents so that they can make a proper route choice decision. Hence, control center operators, aside from devoting their full efforts to clear incidents, also need to decide the most cost-effective management strategy in response to the estimated traffic impacts, such as only offering travel time information of alternate routes to roadway users, or executing either advisory or mandatory detour operations. To do so in an efficient and reliable operating process would demand the availability of a reliable prediction model that can provide the best estimate of a detected incident's clearance duration and its resulting impacts, based on the time-varying information available in real-time incident response process. A robust and efficient model for such a task, however, is not available for use in practice and remains a research subject in the traffic community.

1.2. Objectives and scope of research

The objective of this research is to develop a robust support system for incident management that can provide reliable information for estimating the incident duration and capacity drop due to incidents on a roadway segment. The proposed support system consists of two components, one for estimating the incident duration and the other for computing the resulting capacity drop of the roadway segment blocked by the incident. The former component serves to produce the estimated clearance time of detected

incidents, based on the quality and availability of data collected in real-time operations. It also contains a specially-designed supplemental module to tackle the estimation for those incidents with different natures and high variance of clearance time. The latter component is mainly for assessing the additional reduction in the roadway capacity due to the lane blockages by incidents and the response operations. In addition, this study will also demonstrate the potential of employing the proposed system by the traffic operation center in responding to incidents and managing their impacts.

1.3. Organization of the dissertation proposal

Based on the proposed objectives, this study has divided all research tasks into six chapters. Figure 1 illustrates the organization of this dissertation proposal and the logical relations between its principal tasks. A brief description of the key research tasks in each chapter is presented below:

- **Chapter 2** summarizes the results of a comprehensive literature review associated with the estimations of incident duration and capacity drop which are the key components in an effective incident management system. Also included are the additional research needs along with the focus of the study.
- **Chapter 3** illustrates the overall structure of the proposed incident management support system. Key components of the proposed system and their interactions in the system's operating process constitute the core of the chapter. Also included are descriptions of the inputs and outputs of the system, and the information produced

from each principal component along with their roles in the incident management process.

- **Chapter 4** first presents the nature of the incident duration data along with detailed descriptions of the data pre-processing tasks prior to the model development. This is followed by a presentation of the classification model for incident duration to provide the initial interval estimate with available information from the incident scene. This proposed classification model will be used to yield an initial categorization of detected incidents using the information of the incident types and the number of blocked lanes, and then applying the mining method with hybrid association rules to achieve a more robust estimation for each subcategory of incidents.
- **Chapter 5** details a continuous model to provide the point estimate of a detected incident's duration which is essential for assessing the spatial and temporal incident impacts. The proposed continuous model will be used to divide the incident duration data into several subgroups, based on those factors exhibiting significant impacts on the resulting long clearance durations. This chapter will also discuss the algorithm to identify the most appropriate modeling techniques to each subgroup, and to integrate the results with the estimates from the primary model to yield a robust estimate.
- **Chapter 6** describes the algorithm and the process to distinguish the data anomalies from faulty noises in the incident dataset. A detailed illustration of using the information from those identified outliers to enhance the accuracy of the primary prediction model for incident clearance time constitutes the core of the chapter. Some numerical results to assess the contributions of the proposed outlier analysis algorithm have also been presented in this chapter.

- **Chapter 7** introduces the key idea and mechanism for computing the additional capacity reduction (i.e., capacity drop) due to the lane closure by a detected incident. By using the parameters that significantly affect the capacity drop at the lane-closure roadway segment (e.g., vehicle speed, acceleration rate, merging time, and merging ratio), this chapter has developed an efficient analytical model for estimating the capacity drop under a single-lane-open scenario with homogeneous flows by using the concepts of effective voids and merging spaces. Also, a set of numerical analyses has also been conducted to investigate the model's properties.
- **Chapter 8** extends the single-lane-open model for estimating the capacity drop in Chapter 7 to a multi-lane-open model, considering the heterogeneities in traffic flow properties. The merging maneuvers under those generalized conditions can be divided into mandatory and discretionary lane changes. The effective voids and effective merging spaces from the middle-open and right/leftmost-open lanes are used to compute the total capacity drop in the multi-lane-open scenarios. The proposed capacity drop model in this chapter has also been analyzed to investigate its properties using the real incident data collected from the RITIS and its CCTV.
- **Chapter 9** presents an efficient queue and travel delay analysis process in real time, based on the shockwave queuing concept, to deal with the time-varying incident impacts on the highway networks during each phase of the incident management process and its traffic conditions. Also, this chapter has illustrated the proposed system's application with an example that integrates the estimates from the principal components developed in this study to provide the robust estimates of incident duration and traffic impacts to the highway agencies for their decision-making

processes in the real-world incident response and management. This chapter also includes conclusions, contributions, and future research of this dissertation.

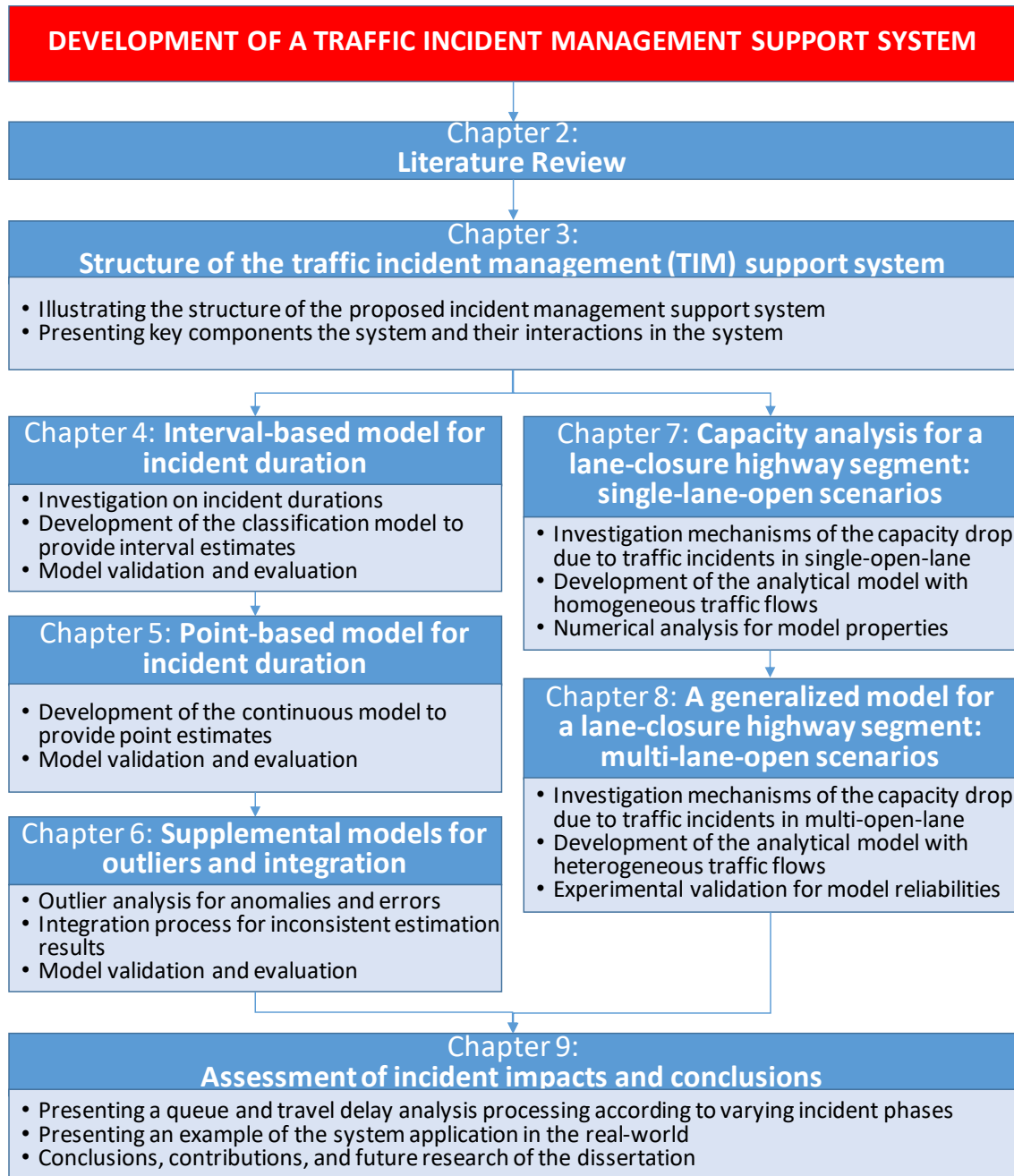


Figure 1. Dissertation proposal organization

Chapter 2. Literature review

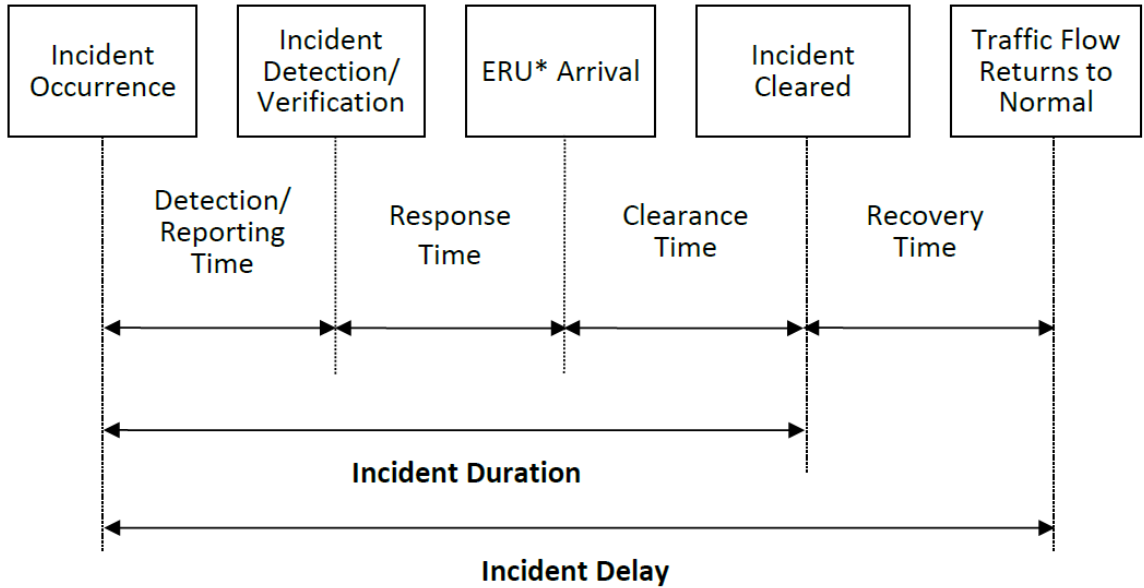
This chapter will first present critical issues associated with incident duration prediction and capacity drop estimation by traffic incidents, and then summarize the results of a comprehensive review of major studies in the literature, focusing on modeling approaches, and potential research directions.

2.1. Estimation of incident duration

Over the past decades, traffic researchers have devoted considerable efforts on the subject of providing a reliable estimate of the duration needed to clear a detected incident, and explored a variety of methods to capture the complex compounding impacts of various factors on the resulting incident duration. Examples of such estimation methods include statistical models, classification/tree-based models, artificial intelligence (AI), machine learning (ML), and hybrid models.

2.1.1. Definition of incident duration

According to the Highway Capacity Manual (HCM) 2010, one can divide the entire incident duration into four phases: 1) incident detection time (time duration from the onset of an incident to it being detected and reported), 2) response time (the incident detection time to the arrival of the response team to the scene), 3) clearance time (from the arrival of the response team to completion of all clearance works), and 4) recovery time (from the incident clearance time to full recovering of the traffic condition).



*ERU: Emergency Response Unit

Figure 2. Components of traffic incident duration [4]

Generally, incident duration can be defined as the time difference between the onset of an incident and the end of its clearance time, even though some studies also include the recovery time as the incident duration. Hojati et al. [5] used the total duration including the detection and recovery times in their study, Kaabi et al. [6] developed a model to predict the incident response time. Also, Nam & Mannering [7] and Li [8] focused on different time phases in an incident response and management process such as detection, response, and clearance time of the incident. They showed that the required duration for each stage of incident response and management may vary with different factors.

Nevertheless, most studies in developing an incident duration model employ the incident clearance time between the arrival time of a response unit and the clearance time of a detected incident due to the following reasons: 1) the time duration between the occurrence and detection of an incident is not always available, 2) the response time is

more related to available person-power of the incident management program and the traffic conditions, but not associated with the nature of incident and the clearance efficiency. 3) the recovery time is difficult to measure precisely, and 4) most incident response teams use the incident clearance time as measures of performance evaluation. Hence, in view of such data limitation and the nature of the data recording system, most studies use the incident clearance time as the estimator in their incident duration models [9].

2.1.2. Nature of incident duration dataset

Note that Incident duration data, as reported in the literature, have unique distribution patterns and complex correlations among contributing factors.

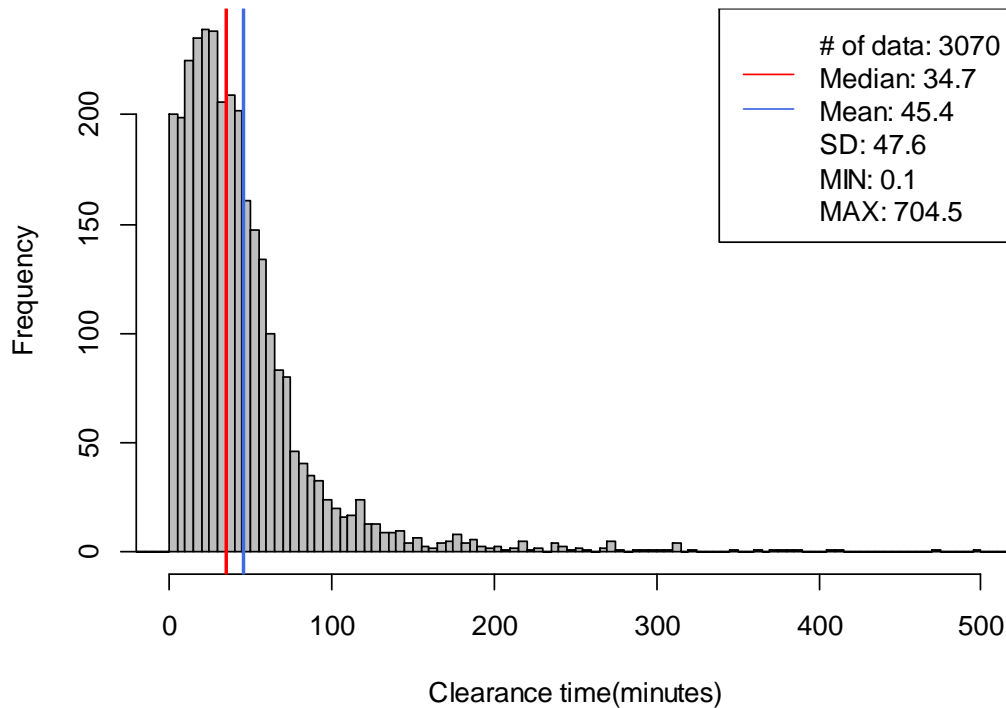


Figure 3. Distribution of the incident clearance time for MD I-95 in 2012-2017 (Collision only) [10]

As shown in Figure 3 incident duration data generally have a right-skewed distribution, but those distributions can also vary with the data sources and their incident types [10]. For example, some studies confirmed that the incident duration data may follow the log-normal distribution [11-13], log-logistic distribution [7, 14-18], or Weibull distribution [5-7, 15, 19]. Other researches, such as Ghosh et al. [20], claimed that the incident duration follows the generalized F-distribution. Furthermore, some of the studies revealed that the resulting distribution patterns actually vary with incident nature, and the surrounding environmental conditions [5, 7, 8, 15, 21]. Thus, selection of the appropriate distribution assumption is one of the key tasks in developing a reliable incident duration model [9]. A recent study by Zou et al. [22] claimed that the mixture models might be a potential direction for tackling the distribution issue in developing incident duration models.

Won et al. [10] claimed that the resulting incident duration depends on a large number of factors and associated variables such as the nature and location of the detected incident, weather and environmental conditions, resources and staff level of the incident response team, and incident onset time and the traffic volume. Also, the incident records for model development often suffer from extensive faulty data or missing critical variables (e.g., blocked lanes, number of injuries, arrival and departure time of the response team, and tow trucks) due understandably to the emergency nature of incident response and the priority of involving staffs. Lastly, many of their contributing variables are either quality in nature or in need of being documented at the acceptable level of accuracy (e.g., opening time for each blocked lane, arrival time of the response unit).

2.1.3. Statistical approaches

In estimating the incident duration, many studies also have employed various statistical modeling methods ranging from a conventional regression to an advanced one [7, 11, 12, 17, 18, 21-34]. Garib et al. [25] developed a regression model for the incident duration using around 200 incident cases. From this study, they have identified some significant contributing factors to incident duration, such as that the number of lanes, the number of involved vehicles, truck involvement, time of day, police response time, and weather conditions. Peeta et al. [35] developed two different models according to the incident types, such as crashes and debris, and used the predicted results for providing the real-time traffic advisory and route guidance. Khattak et al. [29] produced a similar model with the ordinary least squares (OLS) regression using around 16,000 incident cases. They reported to achieve approximately 37% of the overall MAPE for incidents with incident duration from 10 to 120 minutes. Lastly, Yu & Xia [36] first divided the collected incident data into traffic accident and vehicle assistance, and then developed the regression model to predict the duration.

For the same purpose but with different methods, several researches applied the hazard-based duration model to deal with the right-skewed distribution pattern and the probability among the contributing factors to impact the incident duration [7, 18, 21, 22, 27, 28, 31, 33]. Among those, Nam & Mannering [7] employed the hazard-based model to statistically analyze the detection/reporting, response, and clearance time in the incident duration, using the 2-years data (681 incidents) from Washington State's incident response team program, and the Weibull distribution for the incident detection/reporting and response time, and the log-logistic distribution for the incident clearance time and response

times. They claimed to uncover the same insightful information about the unique characteristics of incident detection and clearance time distributions which cannot be identified with the use of standard regression methods.

Chung [28] presented the log-logic accelerated failure time (AFT) metric model to predict the incident duration, based on an accident dataset (2,940 accidents from years 2006 to 2007) from the Korean Freeway Systems. The dataset of the year 2006 was utilized to calibrate the model, and the dataset of the year 2007 was employed to test the temporal transferability of the calibrated model. Their study results indicate the existence of critical relations between the incident duration and its key associated factors, including the number of involved vehicles, the number of injured people, fatality, types of involved vehicles, accident types, location, and report types.

Similarly, Junhua et al. [18] developed a log-logistic AFT model to conclude the impacts of the following factors on incident duration using Chinese policeman's freeway incident record of 1,198 accidents: reporting period, time for the policeman to reach the site, incident type, lanes blocked, vehicle type, vehicles involved, the number of rescue vehicles, and the number of fatalities. In addition, they claimed that using the AFT model can better circumvent the issue of missing data in model development and calibration.

Hojati et al. [21] also examined the impacts of numerous factors associated with the resulting duration of various incident types, using one year's incident data from the Australian freeway network (4,926 accidents). Variables identified to significantly impact the incident duration include the characteristics of the incidents (e.g., severity, type, towing requirements, etc.), location, time of day, and traffic characteristics of the incident. Infrastructure and weather were not significant in their study.

Li and Shang [31] employed not only several parametric AFT models, but also flexible parametric AFT model to examine the effect of numerous factors associated with various types of incidents. A flexible mathematical function is defined by the piecewise polynomial. Due to the diversified distribution of traffic incident duration data, the general AFT model may not be flexible enough to adequately represent the hazard function and capture the underlying shape of the incident duration data [37, 38]. Similar to many previous studies, this study also found that the best distribution for different phases of the incident duration may vary with incident nature and many other factors. Noteworthy, although semiparametric models may offer a better flexibility to fit the distribution of incident duration data, such model models cannot yield the anticipated performance, and often produce large errors for incidents of long duration.

With the same focus, Zou et al. [22] also employed the finite mixture model, an alternative approach in the field of survival analysis that offers better flexibility in describing different shapes of the hazard functions. Additionally, the finite mixture model is grounded in the assumption that the dataset for incident clearance time contains distinct subpopulations, and it allows the effects of explanatory variables to vary between different subpopulations [22]. In their study, the g-component mixture model was applied to analyze the incident clearance duration with the dataset from freeway segments in Seattle. The results in their study showed that the proposed mixture model can better describe the survival and hazard probabilities of incident duration, and provide more accurate prediction than with the AFT model. Also, the authors claimed that the mixture modeling approach is best suited for analyzing the heterogeneous properties in the incident duration data.

Some studies also utilized the quantile regression method to effectively deal with the unique distribution pattern of the incident duration data, and reported to achieve a more accurate estimation with respect to the incidents of a long duration on the tail of the right-skewed distribution [32, 34]. Examples of such studies can be found from the study by Khattak et al. [32] who developed a quantile regression model to enhance the accuracy of the skewed distribution of the incident duration with more than 85,000 incident cases from the years 2013 to 2015 in Hampton Roads area of Virginia. The collected incident data were divided into ten subgroups according to their quantile values. The results indicated that the estimated parameters and the prediction results were quite different, compared with those from the ordinary least squares (OLS) regression model. Those incidents of longer duration can be estimated more accurately. Then, the developed model could also estimate the probability of a detected incident to last over a specified duration.

Zou et al. [34] also employed the quantile regression model to analyze the effects of explanatory variables on the resulting duration. Using the incident data collected from the freeway sections in Seattle, the authors claimed that the quantile regression can be effective for estimating the effects of explanatory variables on the clearance time distribution.

2.1.4. Classification/Tree-based approaches

Due to the difficulties in identifying proper distributions for incident data, some researchers have attempted to use the classification or tree-structure methods to develop a reliable incident duration model [39-46]. Those rule-based approaches can circumvent the needs to rely on some assumptions about the underlying data distribution as well as correlations

among the contributing factors, offering a flexible way for efficiently capturing abnormal patterns on the tails of the right-skewed distribution.

Along the line, Ozbay & Kachroo [39] employed the decision tree model to predict the incident duration, using the incident data from the Northern Virginia region. In their study, the linear regression technique was first applied, and the hybrid decision tree, similar to the classification and regression tree (CART), was then developed for those incidents not following either the lognormal or log-logistic distributions. Smith & Smith [46] utilized the classification and regression tree (CART) algorithm to develop the decision tree for the incident duration. They used 6,828 collected incidents from the Smart Travel Lab in Virginia and separated the incident duration into three categories, such as '<15' minutes, '15-30' minutes and '≥ 30' minutes. They also found the following five significant factors for use of the splitter in developing the classification tree: TOW response, EMS response, police response, day of week, and more than two vehicles involved.

Ozbay & Noyan [40] also utilized the Bayesian Networks (BNs) to develop a dynamic incident duration estimation tree using the incident data collected from the Northern Virginia (around 700 incidents), which may contain partially incomplete information. They demonstrated the advantages of using the BNs in modeling and analyzing the incident duration data because of their abilities to consider the stochastic nature of the data. Unlike the conventional tree models, the BNs utilize the probabilities of each contributing factor to produce stochastic nodes. Thus, the developed model in this study could help the operators estimate the incident duration, which may be concurrently affected by several contributing factors.

Similarly, Boyles et al. [41] employed the naïve Bayesian classifier, which can accommodate incomplete information received at different points in time. Such models, calibrated with the incident data from the Georgia Department of Transportation, showed a better prediction performance than with the standard linear regression model. Zhan et al. [45] utilized the M5P tree algorithm, which can concurrently deal with both categorical and continuous variables, and also handle variables with missing information. One main difference between the CART algorithm and the M5P algorithm is that the regression tree constructed by the CART can have one available value at their leaves, whereas the regression tree built by the M5P can have a multivariate linear model at their leaves. Their developed model showed that the resulting performance is superior to those with the regression and decision tree by the CART algorithm.

With a similar logic, Ma et al. [44] proposed the method with the gradient boosting decisions tree (GBDT) to tackle the nonlinearity and imbalanced distribution in the incident dataset, based on different types of explanatory variables. They used one-year incident data from Washington state and divided the incident duration into two categories: short (< 15 minutes) and long (≥ 15 minutes) durations. The conclusion was that the GBDT method can significantly outperform the most existing algorithms, such as random forest (RF), artificial neural network (ANN), support vector machine (SVM), for incidents with both very short and very long duration.

In summary, the classification/tree-based models are notably straightforward and clear to understand the modeling results. They can overcome the various assumptions needed for using the statistical models, and also efficiently reflect the unique nature of the incident

duration data. However, such models may suffer from the overfitting and transferability issues, and thus limit their potential for applications.

2.1.5. Artificial intelligence (AI) and Machine learning (ML) approaches

With the advances in computing technologies, various artificial intelligence (AI) and machine learning (ML) algorithms have shown their potential for tackling various complex data related issues. Hence, several studies also have applied various such algorithms to the incident duration data with attempt to produce a robust prediction model [47-54].

On this regard, Wang et al. [47] developed the incident duration models with the fuzzy logic (FL) and the artificial neural network (ANN) to compare their resulting performance. They used 695 incident data cases from vehicle breakdowns on the Road Network Master Database (RNMD) in the U.K. for model development, and concluded that both modeling methods have difficulty in predicting the outliers.

Wei & Lee [48] produced the sequential forecasting system for the incident duration using the ANN approach. Their study consists of two models, where the first was used to predict the incident duration at the time of incident notification, and the second was to provide multi-period updates of the incident duration in the following period. The study, based on 24 incident cases from the National freeway in Taiwan, classified the various types of data into incident characteristics (e.g., traffic data, time and space gap) and geometric characteristics. The evaluation results showed that their developed model can well fit the actual incident duration, and the ANN model can effectively smooth the data noise.

Wu et al. [51] focused on developing an incident duration model with the support vector regression (SVR) which can effectively tackle the issues associated with small sample size, nonlinearity, and high dimensional features. Using the dataset from the Netherlands Ministry of Communication, the authors claimed that their developed model can yield the reasonable level of accuracy and adaptability in the duration prediction.

Park et al. [53] utilized the Bayesian ANN to predict the incident duration, and the pedagogical rule extraction algorithm (TREPAN) to extract comprehensible rules from the calibrated neural networks. They used 13,987 incidents collected from the Maryland Highway Administration (SHA) and compared their estimated model with other prediction techniques in the literature.

2.1.6. Combined/Hybrid approaches

Some studies have connected different modeling methods and develop a hybrid model to produce a more robust result [10, 55-62]. The logic behind such methodology is that different modeling methods with different assumptions may perform better with different parts of the data. Thus, a proper combination of the estimates from these models may provide more robust results than using any of their individual components [63].

Lin et al. [55] integrated the discrete choice model with the rule-based method to estimate the duration of a detected incident, using the data from the Maryland State Highway Administration. They applied the ordered probit model to the incident cases with less than 60-minute duration, and then used the rule-based model as a supplemental module to the incidents exceeding 60 minutes to capture complex interactions among the contributing factors to the incident duration. From such integrated modeling process, the

proposed model was reported to achieve more than 83.1% of the overall prediction accuracy and provided a reliable estimate for severe incidents that may last more than one hour.

In the studies by Kim et al. [56] and Kim & Chang [57], they developed the similar hybrid models which are composed of the primary model and supplemental models to enhance unsatisfactory estimates from the primary model. Kim & Chang [57] utilized the rule-based tree model (RBTM) as the primary model, and applied the multinomial logit model (MNL) or the naïve Bayesian classifier (NBC) to the nodes with an unsatisfactory result as the supplemental models. The incident dataset from the Maryland State Highway (MDSHA) and Police Accident Report database between years 2003 and 2005 were used for model development. The developed model performed satisfactorily for most incidents, except for those in the categories of ‘60-120’ minutes and ‘240-300’ minutes.

Ji [58] proposed to combine two estimates from the regression model and the decision tree model using their multi-model fusion algorithm to enhance the prediction performance. The idea of their fusion algorithm is to first calculate the estimation errors from the regression and decision tree models, and then to compute the weight for each model to combine them. He et al. [59] proposed the hybrid tree-based quantile regression which incorporates the following strengths from both modeling methods: robustness to outliers, simple for interpretation, flexibility in combining categorical covariates, and capturing nonlinear associations. They used the incident-related data as well as the traffic-related data before and after an incident occurrence for their model development. In the modeling step, they first developed the decision tree using the unbiased recursive partitioning (URP) algorithm, and then applied the quantile regression method to each terminal node in URP

to achieve a better prediction accuracy. In the evaluation, they claimed that the hybrid tree-based quantile regression model in their study provided a more accurate prediction result than with other models such as the classification and regression tree (CART) and the K-nearest-neighbor (KNN).

Lin et al. [61] explored the method of combining the M5P tree and hazard-based duration model (HBDM) to produce the new M5P-HBDM algorithm for predicting the incident duration. The M5P algorithm, using the multiple linear regression models as its leaves, can mostly outperform those the traditional classification and regression tree (CART) algorithms. However, it has the weakness of assuming the normal distribution in developing the multiple linear regression for each node. By using the HBDM in each node, instead of the multiple linear regression model, to reflect the unique distribution of the incident duration data, they claimed to achieve acceptable performance in predicting incident duration, using 602 incident duration data in I-64 Virginia from the Virginia Department of Transportation (VDOT's) Archived Data Management System (ADMS).

2.1.7. Outlier analysis

As shown in Figure 3, the distribution of the incident duration data is highly right-skewed, and exhibits a long tail on the right side of the distribution. Won et al. [10] indicated that about 80% of incident clearance times in most datasets are less than 60 minutes, and only less than 10% exceed 120 minutes. Also, the maximum incident clearance time is approximately 12 hours [10]. This means that a small portion of such cases shows a pattern quite different from most cases, but they can have a massive impact on the model

calibration and development, especially when the data have a long tail and unspecified distribution.

Li [8] analyzed the incident duration and developed the prediction model using the hazard-based model to consider the unobserved heterogeneity, time-varying covariate, and relationship between consecutive traffic incident duration stages. Their evaluation result showed that the developed prediction model performs well for most incidents in a normal range, but it does not have the acceptable performance for incidents with a very short (less than 15 minutes) or long (more than 120 minutes) duration.

Khattak et al. [29] developed an incident duration prediction model, using the ordinary least squares (OLS) regression model with almost sixty thousand incident cases. Their statistical model has around 37% of MAPE, but yield unreasonable estimation results for the incidents with a very short (less than 10 minutes) or long duration (more than 120 minutes).

Valenti et al. [49] applied different prediction models to the same incident duration dataset to investigate their reliability. They used five different modeling approaches, ranging from the statistical model to the neural networks, such as multiple linear regression (MLN), decision tree (DT), artificial neural network (ANN), support vector machine (SVM) and K-nearest-neighbor (kNN). Their evaluation results showed that all developed models show a good performance in terms of prediction accuracy, especially for incidents lasting less than 90 minutes. Only the ANN model has showed to achieve the reliable prediction results for incidents exceeding 90 minutes.

In summary, many other studies also have reported to realize such outlier issues, but no effective model or methodology has been developed. Because it is an unsupervised

learning problem and a highly imbalanced classification problem, and most statistical methods and convergence algorithms tend to capture the central tendency in the data rather than the outliers to a certain extent [9].

2.2. Estimation of capacity reduction

Many previous studies have tackled to estimate capacity drops at the lane drops such as work zones, ramps, traffic incidents, etc. There are several approaches, and they can be categorized into four groups: parametric approaches [64-71], non-parametric approaches [72-76], simulation-based approaches [77-80], and analytical approaches [81-94].

2.2.1. Definition of capacity reduction

Highway incidents usually close one or more of the available travel lanes and disturb the normal traffic flow. Those lane closures may increase the number of traffic conflicts, and the outflow from the lane drop area is the same as the queue discharging rate [95]. Those available flows are generally lower than the free-flow capacity, and this phenomenon is called the capacity drop or additional capacity reduction [96].

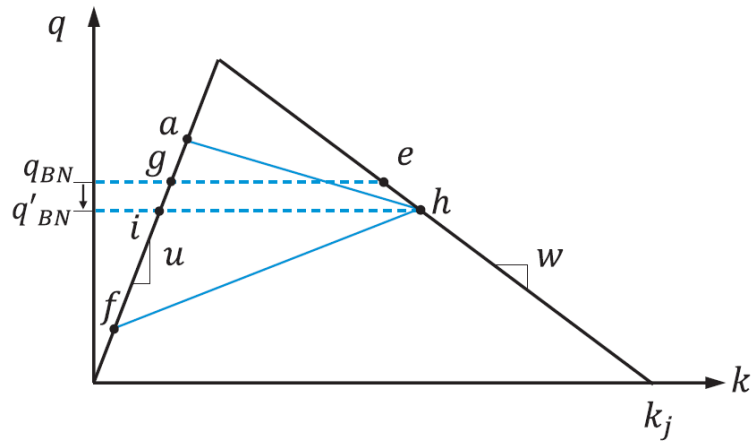


Figure 4. Example of the capacity drop in the fundamental diagram [97]

Figure 4 shows an example of the capacity drop phenomenon in the triangular fundamental diagram with free-flow speed u , wave speed w , jam density k_j , demand a

[97]. If the bottleneck by an incident is operated as its capacity without traffic breakdown, the outflow, q_{BN} , should be the same as the maximum capacity of the opened lanes with traffic states in the upstream, e , and downstream, g . For example, if one out of two lanes are blocked, the outflow, q_{BN} , would be the half of the maximum capacity of the two lanes. However, after the onset of queue and breakdown, the actual discharge rate, q'_{BN} , may be lower than the bottleneck capacity, q_{BN} , due to its traffic conflicts. That additional capacity-reduction phenomenon represents the capacity drop by an incident in this study.

In many empirical studies, those capacity drops are commonly observed between around 5% and 30%. Hall & Agyemang-Duah [98] investigated the capacity drop when a queue forms using the data on the Queen Elizabeth Way (QEW) in Canada. They concluded that there is a capacity drop in the bottleneck. Once a queue has formed upstream, the bottleneck location does not handle as many vehicles as it did before the queue formation. Those capacity drops were about 5~6%. Banks [99] observed the capacity drop in four freeway bottlenecks in San Diego. The flow immediately downstream of the bottlenecks decreased by around 3% when it broke down, and that flow breakdown appeared to be triggered by speed instability. Kerner [100] investigated congested patterns measured during the year 1995-2001 at German highway. He found the synchronized flow pattern at an isolated bottleneck and around 12% capacity drop. Cassidy & Bertini [101] observed the capacity drop phenomenon from two freeway bottlenecks in and near Toronto in Canada. The average rate vehicles discharge from a queue can be 10% lower than the flow measured prior to the queue's formation. They also found that the queue discharge flows exhibited nearly stationary patterns that alternated between higher and lower rates.

Some studies tried to investigate various relations between the capacity drop and related traffic conditions. Sarvi et al. [102] investigated traffic behaviors and characteristics during freeway ramp merging under congested traffic condition on the Tokyo Metropolitan Expressway. In their study, the capacity of the merging sections was about 7~18% lower than an ordinary highway section, but this capacity was not correlated with the merging ratio when both the ramp and freeway approaches are queued. Chung et al. [103] investigated relation between vehicle density and losses in discharge flow in different geometry conditions. They considered three different bottlenecks such as an on-ramp merge, a reduction in travel lanes, and a horizontal curve. The finding of this study is that those different geometry conditions and densities showed 3~18% capacity drops in the field data.

Several studies also focused on when the traffic flow breaks down and how the capacity drop is triggered. Cassidy & Rudjanakanoknad [104] verified that discharge flow diminishes when a merging becomes an active bottleneck using the video camera data on Freeway 805 in San Diego, California. They found that traffic breakdowns always occurred at or shortly before the times measured occupancies rose to 27%, and the deleterious shoulder lane accumulations were always reduced below 16 vehicles at or shortly before the time occupancies dropped to 22%. Bertini & Leal [105] studied traffic conditions in upstream and downstream of the bottleneck that arose near a freeway lane drop near London, the U.K. using archived high-resolution loop detector data. They found that the bottleneck's discharge flow was about 10% lower than the prevailing flow observed prior to queue formation. During the bottleneck discharge, oscillations arose in the queue and propagated upstream at nearly constant speeds of 11~12 mile/hours. Flows measured at

locations downstream of the bottleneck were not affected by these oscillations. Srivastava & Geroliminis [106] showed empirical evidence from freeway-ramp merges in the Twin Cities freeway system and level of the capacity drop depending on the ratio between mainline and ramp flow. They told that the throughput stagnates at a capacity of about 15% decrease from the initial peak when the breakdown happens, and the capacity drop is consistent across time and variation in demands of the mainline and ramp. Oh & Yeo [107] analyzed individual vehicle trajectories at a microscopic level using NGSIM data of US 101, and they found discharge rate reductions near active bottlenecks. Those capacity reductions were observed as ranging between 3% and 18%.

2.2.2. Parametric approaches

The parametric approaches are based on the field data to estimate capacity drops. An appropriate mathematical form is predetermined, and the coefficients of this predetermined mathematical form are estimated by using capacity data collected from the fields. Several studies have developed regression models using the field data from the lane drops in work zones. Krammes & Lopez [64] developed a parametric model to estimate the capacities of short-term freeway work zone lane closures using the field data in 33 work zones in Texas. From those data, they found that average capacities for two-of-three-lanes and one-of-two-lanes blocked conditions are significantly higher than older values reported in the Highway Capacity Manual (HCM) 1985, and work intensity, presence of ramps, and heavy vehicles have significant impacts on the capacity of the short-term freeway work zones. Nevertheless, this parametric model lacked to deal with various contributing factors on the capacity and their complicated relations. Kim et al. [66] tried to take into account various

contributing factors on the capacity reduction in work zones and developed a multiple regression model. Geometric data were collected at 12 work zone sites with lane closures on four normal lanes in one direction, mainly after the peak-hour during daylight and night. And then, they found several key factors of the developed capacity-reduction model, such as the number of closed lanes, lane closure locations, heavy vehicle percentage, lateral distance to the lane closure, work zone length, work intensity, and work zone grad. Lastly, they compared their developed model with other existing capacity models and showed improved performance for all of the validation data. Al-Kaisy et al. [65] investigated freeway capacity at long-term lane closures due to rehabilitation work in Ontario, Canada. The field data collected from two-of-three-lanes closures work zone, at the same site with eastbound and westbound, for four days. They found significant temporal variations in freeway capacity in the work zone, but the average capacity is very close to the corresponding values provided in the Highway Capacity Manual (HCM) 1997. They also developed a multiple regression model to estimate long-term work zone capacity and the significant factors in their proposed model were the temporal variations, day of the week, weather conditions, and grades. Al-Kaisy & Hall [67] investigated the capacity at freeway reconstruction sites in Ontario, Canada, and analyzed the effects of several contributing factors to the capacity of lane closures, such as effects of heavy vehicles, driver population, rain, site configuration, work activity, and light condition. And then, they first developed a site-specific capacity model using the data from the site where has the most extensive and comprehensive capacity data in their study. Based on this developed site-specific model, they extended it to a generic capacity model using six work zone sites. The proposed model in their study showed that a base capacity value of 2,000 passenger cars per hour

per lane for reconstruction sites under favorable conditions, and heavy vehicles and driver population were the most significant factors on the capacity reduction.

Benekohal et al. [68] argued that most statistical models strongly depend on the quality of the collected data and how to measure the capacity from the field is a critical issue to guarantee the model reliability. Thus, they proposed an alternative method using the speed-flow relationship to estimate the capacity drop due to lane closures. The idea of their study was based on that work intensity, lane width, lateral clearance, and other operating factors cause drivers to reduce the speed, and those speed reductions will affect the capacity drop in work zones. Video data were collected for around 30 hours from 11 work zones on highways in Illinois where one of two lanes was closed, including long-term and short-term work zones. Using those collected data, they analyzed the relations with the speed reduction and developed a speed-flow curve to estimate the work zone capacity. Sarasua et al. [69] investigated various contributing factors on the capacity of short-term lane closures for Interstate work zones in South Carolina, such as roadway grades, work zone activity, weather conditions, etc. They analyzed the traffic characteristics of freeway lane closures using classic methods of macroscopic traffic flow modeling. Relationships of speed, density, flow, and time headway were derived based on the linear relationship between density and speed in the traffic flow theory. And then, they developed a model for calculating the work zone capacity that incorporates base capacity, passenger car equivalents for various speed groups, adjustment factors related to specific work zone characteristics, and the number of lanes open through the work zone. Racha et al. [70] developed a non-linear hyperbolic model of work zone traffic flow elements to describe the relations between speed, flow, and density that can be used to estimate the capacity in

work zones. Traffic flow data in work zones were collected from 22 sites on South Carolina interstate highways. In their study, the work zone capacity was estimated to be 1,550 passenger vehicles per hour for one-of-two-lane closures, and adjustment factors for the capacity, such as types of vehicles and work zone intensity, were estimated. Avrenli et al. [71] examined the speed-flow relationship of a selected work zone with lane closure. They estimated the speed-flow curve as a three-regime model that involved free flow, transition, and congestion regimens. The average capacity of the lane closure was estimated to be 1,900 passenger car per hour per lane, and they also proposed the step-by-step procedure to predict the capacity of the lane closure.

2.2.3. Non-parametric approaches

The parametric approaches have difficulties in dealing with a lack of data points of the collected field data and describing the complicated nonlinearity and hidden intercorrelation among influencing factors. Hence, many studies have also applied the non-parametric approaches to estimate capacity drops. Adeli & Jiang [72] employed a neural network to estimate the parameters associated with the bell-shaped Gaussian membership functions used in the fuzzy inference mechanism, which included various contributing factors such as percentage of truck, pavement grade, number of lanes, lane width, work intensity, length of closure, lane closure duration, weather and pavement conditions, driver composition, etc. The developed model in their study showed a more accurate estimate than others in the literature, especially when the data for key contributing factors to the capacity are only partially available. Also, their model could incorporate many factors impacting the capacity reduction. Karim & Adeli [73] developed a radial-basis-function neural network model to

describe the work zone traffic condition associated with the capacity, using the information about the number of lanes, work zone layout, percentage trucks, speed, darkness, proximity of ramps, etc. The developed model in their study exhibited good generalization properties from a small set of training data, an especially attractive feature for estimating the work zone capacity where only limited data is available. Zheng et al. [75] gave a comparison study of the existing models such as Highway Capacity Manual (HCM) 2000, two multi-linear regression models, and a fuzzy logic based artificial neural network model, using Dutch freeways. The comparison results showed that the neuro-fuzzy model had the highest average accuracy, and the prediction error could be reduced as large as 20%. But they claimed that those artificial intelligence models are too complex to understand and inappropriate for application in the field. Weng & Meng [74] developed a decision tree-based model for estimating the freeway work zone capacity using the collected data from 14 states and 16 influencing factors, such as percentage of heavy vehicle, work zone grade and intensity, road type, work zone duration, weather, driver composition, ramp, speed, state, etc. Their statistical comparison results showed that the decision tree-based model in their study outperformed the existing short-term and long-term freeway work zone capacity estimation models in the literature, especially when the input values of influencing factors are only partially available for the existing models. Weng & Meng [76] also proposed a bootstrap aggregation method that is used to build an ensemble tree comprising a set of individual decision trees. The results of the statistical comparison showed that the ensemble tree performed better than any single decision tree and reduced the instability that is one of the shortcomings of the tree-based models.

2.2.4. Simulation-based approaches

In the real world, collecting enough qualitative field data related to the capacity is really challenging. Hence, many studies have also employed simulation-based approaches to acquire enough data points or directly estimate the capacity at the lane closure areas. Chatterjee et al. [78] tried to estimate the traffic impacts of the work zones using microscopic simulation program, VISSIM. They had noticed that the capacity of lane closures might vary from state to state; thus, the default parameter values in the simulation models are not suitable for the capacity analysis in the lane closures. Hence, they proposed a simple method for choosing appropriate values of driving behavior parameters in the VISSIM microsimulation model to match the desired field capacity for the work zones operating in a typical early-merge system. Two most significant car-following parameters (i.e., desired time headway and longitudinal following threshold) and one lane-changing parameter (i.e., safety distance reduction factor) were selected to estimate different lane closure environments. Heaslip et al. [79] developed analytical models and procedures for estimating the capacity of freeway work zones using simulation-generated data and field data. First, they used CORSIM to generate a comprehensive database for various work zone scenarios and developed analytical models to predict the capacity of lane closures. And then, those developed models were evaluated and refined by using the collected field data from I-95 in Florida. The developed models in their study showed within 1% error between the estimates and actual capacities in their study. Heaslip et al. [80] also utilized the simulation method in the previous study [79] to estimate the capacity of short-term lane closures on arterial roads. Through their simulation analyses, they concluded that 1) the distance of the lane closure to the downstream intersection can affect the capacity drop of

entire arterial work zone, 2) one of the movements blocked by the others can reduce the capacity, and 3) installing a work zone notification may increase the capacity when the arterial road is congested.

2.2.5. Analytical approaches

Several studies have used analytical approaches to develop a capacity drop model. Those analytical approaches build a mathematical formulation describing mechanisms of the capacity drop phenomenon, based on the relations among the key contributing factors on the capacity drop. Those mathematical formulations are very interpretable and easy to track the core relations for the capacity drops. In the literature, the analytical approaches for the capacity drop have utilized the merging/lane-changing mechanism, gap acceptance theory, etc. These analytical approaches based on the traffic behavioral mechanism can be divided into three categories: inter-driver/vehicle spread [81, 82], intra-driver spread [83, 85, 86, 89], and bounded acceleration capability [90, 91, 93, 94, 108].

First, the idea of the inter-driver/vehicle spread is that each driver and vehicle have a different driving behavior and mechanical performance; thus, conflicts among vehicles can create voids, and these voids can lead to capacity drops. Papageorgiou et al. [81] have claimed that the capacity drop is due to the acceleration difference between two successive vehicles, and voids can be created between a low-acceleration vehicle and its high-acceleration predecessor. Based on those ideas for the capacity drop, they developed a general conceptual framework for real-time merging traffic control to maximize the throughputs and minimize the delays in the lane closure such as merging highways, merging on-ramps, toll plazas, and work zones. Wong & Wong [82] have reproduced the

capacity drop in numerical simulations when formulating a multiclass traffic flow model as an extension of the Lighthill–Whitham–Richards (LWR) model [109, 110] with heterogeneous drivers. The developed model in their study can consider the heterogeneity of drivers characterized by different speed choices in the traffic stream, which a faster vehicle tries to overtake a slower vehicle in the congested and uncongested conditions and a slower vehicle would slow down the faster one. Through the numerical simulations in their study, they confirmed that the developed model could describe some of the confusing traffic flow phenomena such as hysteresis and platoon dispersion in the fundamental diagram.

Second, the idea of the intra-driver spread is that even a driving behavior of a single driver can vary depending on different traffic conditions; thus, these changeable driving behaviors can make the capacity drops. Treiber et al. [83] investigated the adaptation of the time headways in car-following models as a function of the local velocity variance. They assumed that drivers would choose a different time headway according to the different congested conditions, and applied this mechanism to several car-following models in the on-ramp bottleneck. And then, they estimated the distributions of the net time headways for free-flow and congested traffic, the velocity variance as a function of density, and the fundamental diagram. From the analysis results in their study, they could reproduce free, synchronized, and jammed traffic, and a wide scattering in the congested traffic regime. Zhang & Kim [85] presented a multi-phase car-following theory to reproduce both the capacity drop and the traffic hysteresis. The logic behind the proposed model is that the capacity drops and hysteresis, two prominent features of multiphase vehicular traffic flow, are related to drivers' behavioral shifts during phase transitions,

such as acceleration, deceleration, and coasting. Hence, they proposed the introduction of a single variable, such as gap-time, that depends on both vehicle gap-distance and traffic phase. And then, they could reproduce the well-known car-following phenomenon through theoretical analysis and numerical simulation. Yeo [86] proposed the microscopic asymmetric traffic theory based on analysis of individual vehicle trajectories and applied this proposed theory to describe the traffic phenomena in the congested traffic, such as traffic hysteresis, capacity drop, stability, relaxation after lane changes, and stop-and-go waves. The capacity drop was explained as a difference of the maximum flow between the acceleration and the deceleration curve in the density-flow fundamental diagram. Chen et al. [89] showed that traffic hysteresis arises due to variable driver characteristics within each driver and has a profound reproducible impact on the periodicity and development of traffic oscillations and the bottleneck discharge rate. The corresponding capacity drop is a result of the change of aggressiveness described as four different reactions: concave, convex, non-decreasing and constant. They also validated their theory using NGISIM dataset and identified the capacity drop ranging from 8% to 23%.

Lastly, the idea of the bounded acceleration is based on the concepts that vehicles have limited acceleration capabilities and cannot also accelerate instantaneously, thus inserting vehicles would create voids in front of them and it can make capacity drop on the target lane. Several studies have shown that low-speed merging maneuvers in a free-flow stream can reduce available capacity and trigger congestions [90, 91, 93, 108]. Laval & Daganzo [90] hypothesized that lane-changing vehicles might create voids in traffic streams, and those created voids reduce the flow. To be specific, drivers try to change lanes to increase their speed, and those lane-changing vehicles act as a moving bottleneck on the target lane

while accelerating to the speed of the target lane. And then, those ensuing disruptions can trigger other lane changes. Those mechanisms were described by modeling their lane changing decisions to track lane changers. Their analysis results confirmed the capacity drop in flow observed after the onset of congestion at freeway lane-drops and the relations between the speed of moving bottlenecks and their capacities. Also, they concluded that the spatial distribution of lane changes and the difference in lane speeds are important determinants of bottleneck capacity; thus, traffic managers may be able to increase capacity by forbidding lane changes and posting speed advisories at key locations. Duret et al. [91] believed that low-speed merging maneuvers performed within a free-flow stream will trigger traffic congestion by disturbing the flow at a local and global scale. Thus, they first explored the analytical solution of a simple first-order model using the kinematic wave theory when moving boundary conditions are introduced. And then, they proposed an analytical formulation of the capacity drop for multiple merging maneuvers at a single location. Through the developed model in their study, they showed that the capacity drop is related to the demands on the minor and major streams and the speed of the merging vehicle. Leclercq et al. [93] claimed that the Newell-Daganzo merge model is not only very simple but also accurately reproduces experimental findings; however, the capacity of the merge is an exogenous variable in this model. Thus, they proposed an analytical model that extends the Newell-Daganzo model by incorporating the capacity drop related to the merging process and directly estimated quantitative impacts of lane changing, such as the queue discharge rate calculation in the single-lane-with-on-ramp freeway. Their developed model is based on the physical mechanism of bounded acceleration of merging vehicles, and they validated their model using the experimental data coming from a merging on the

M6 freeway in UK. Through their sensitivity analysis and experimental results of the proposed model, they found the following model properties: 1) the low acceleration rate has the highest impact on the capacity drop, 2) the relative capacity drop is decreased when the length of the insertion section (i.e., length of the on-ramp) increases, but too long insertion sections (e.g., longer than 100 meters) do not seem to significantly prevent the capacity drop on freeways, and 3) the merging ratio does not significantly affect the capacity drop when the main road of freeways is only one lane. Leclercq et al. [94] extended their analytical model of the previous study [93] to take into account heterogeneous vehicle characteristics and physical interactions between upstream waves and downstream voids created by inserting vehicles from the on-ramp. Through their extended numerical and sensitivity analysis, they found that accounting for the interactions between the waves and voids can slightly decrease the effective capacity. However, such vehicle heterogeneity is not necessary when only the mean capacity is interested. Hence, they concluded that it is not necessary after all to introduce a detailed description of vehicle characteristics since a proper estimation of their mean characteristics is enough to derive an accurate analytical estimation of the effective capacity. Lastly, Leclercq et al. [108] extended their analytical models of the previous studies [93, 94] again to deal with multi-lane-freeway scenarios with discretionary lane changes that correspond to vehicles that want to avoid the speed reduction. They divided the lane changing area into two nonoverlapping local merging areas such as the mandatory and discretionary lane changes; and then, applied different merging mechanisms to estimate the effective capacities of each open lane in the freeway segment. Their sensitivity analysis of their proposed model showed that vehicle acceleration rates and truck ratios are the most influential factors on

the capacity drop. Those numerical properties of the proposed model were also proven by the results from a traffic simulator that fully describes vehicle dynamics. Finally, they claimed that the proposed model could provide a reliable estimate for the capacity drop in comparison with experimental data for an active merge on the M6 Motorway in the United Kingdom.

Chapter 3. Structure of the traffic incident management (TIM) support system

3.1. Introduction

As mentioned previously, traffic incidents often cause not only a capacity reduction and reliability degradation of a transportation system, but also significant delays to commuters [111]. To mitigate such impacts so as to efficiently recover the transportation system, many highway agencies in the U.S. have established a traffic incident management (TIM) system over the past several decades [3]. A TIM typically consists of a coordinated multi-disciplinary process to detect, respond to, and clear traffic incidents [111]. It is expected that such a system can effectively reduce the clearance duration of detected traffic incidents, and thus reducing their impacts on traffic delays and safety [111].

One essential function of a TIM system is to estimate the required clearance time of a detected incident and its spatial impacts over such duration. As reviewed in Chapter 2, most studies over the past decades have devoted considerable efforts on those vital issues, and produced a vast body of models to provide such a prediction function. However, due to both the data deficiencies and the complex interrelations between key factors contributing to the resulting incident duration and its impacts, most state-of-the-art models on this subject remain at the project demonstration level, but not yet used in practice. In addition, developing a robust system, to deal with the ever-changing incident and traffic conditions caused by sequential incident management processes and implementation of operational strategies, is a critical issue in the real world.

Hence, this chapter presents a TIM system, designed to address all imperative issues reported both in the literature and in the practices of incident responses.

3.2. Key components of a TIM system

A TIM system shall be designed mainly to provide the following functions: 1) reducing the time for incident detection/verification and response, 2) conducting an estimate of incident clearance duration, 3) providing timely and accurate information to the general public for their decision makings [111]. Certainly, the exact system structure and its operating procedures may vary from state to state [111]. Most TIMs, as shown in Figure 5, generally consist of the following eight components for incident detection and impact minimization [112]: detection, verification, response, site management, traffic management, motorist information, clearance, and recovery.

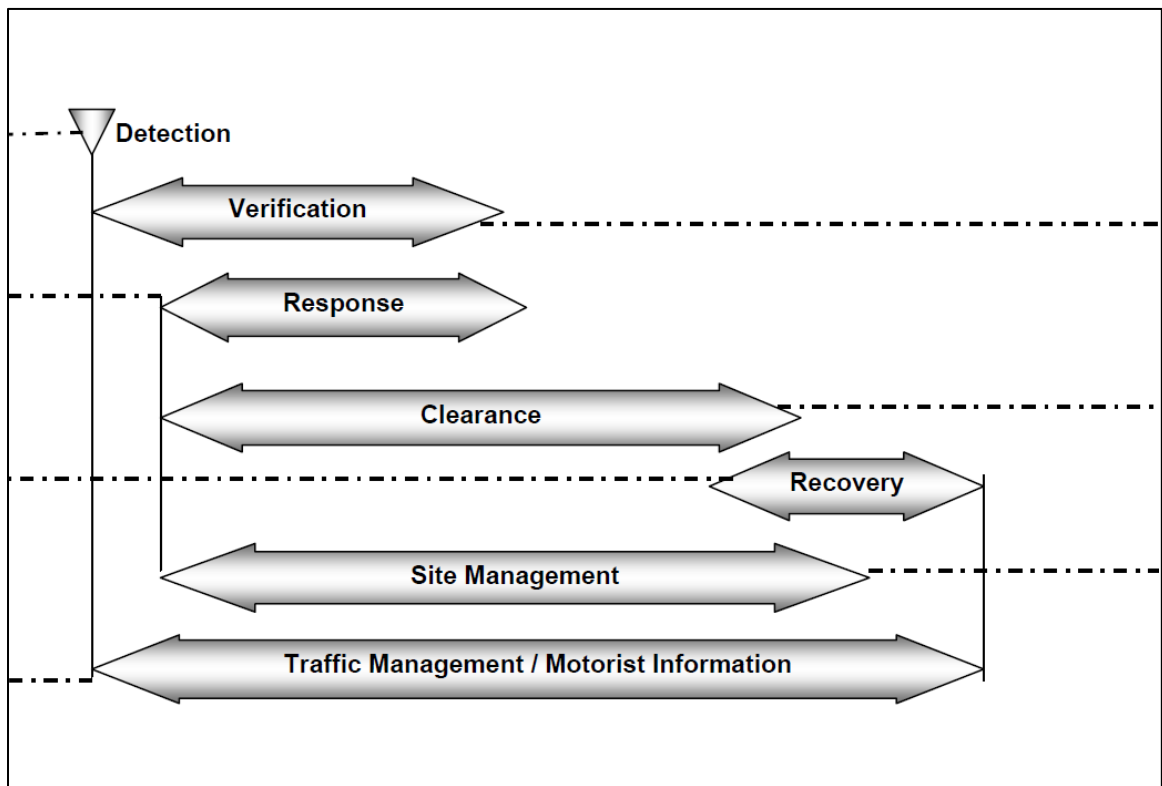


Figure 5. Timeline for the sequence of operations in the TIM process [112]

The first stage in the TIM process is to conduct incident detection and verification. The detection is to identify a reported incident, using various roadside devices and reports from the scene, such as closed-circuit television (CCTV), radar or induction loops with incident detection algorithms, police/service patrols, phone calls, etc. The subsequent task of verification is to further confirm the incident location, incident type, number of blocked lanes and involved vehicles, pavement condition, etc.

The focus of the response stage is to dispatch the appropriate response units to the incident scene, based on the information available from the verification. The level of required response is typically determined by an on-scene responder or by a dispatcher in the communications or traffic management center [113]. A timely response can substantially reduce the duration of the roadway operated under a reduced capacity due to the incident [111].

Site management generally is referred to the tasks of coordination by TMC, which includes accurate assessment of the time-varying incident impacts, establishment of response priorities, activation of an incident command system (ICS) and unified command structure (UCS), notification and coordination with appropriate agencies, and deployment of response vehicles at the incident scenes [111]. Also, traffic management activities during a major incident include activation of detouring operations, allocation of traffic control resources (e.g., cones, warning signs, portable dynamic message signs), timely execution of traffic control strategies (e.g., ramp meters, local traffic signals), and queue management (e.g., monitoring and warning for approaching vehicles) [111].

The task of managing motorist information is to convey incident information to en-route motorists and pre-trip travelers within the incident impact boundaries [113]. Such

information can reduce both the traffic demands and minimize the resulting incident impacts. The methods for conveying the time-varying traffic conditions to roadway users include dynamic message signs (DMS), highway advisory radio (HAR), commercial radio and television broadcasts, telephone information systems (i.e., 511), in-vehicle or personal digital assistant (PDA) information, and information service providers (ISP) [111].

As the final step of the TIM process, its focus is to remove the detected incident in a timely manner from the scene, and safely restore the roadway capacity to its pre-incident level [111]. Note that the recovery task includes the restoration of the roadway to its full capacity, and the required activities generally extend to the debris clearance, restoration of damaged infrastructure, and structural inspections [111].

3.3. Key models to support a TIM system's operations

To minimize the clearance duration of a detected incident and its impacts on the traffic condition, the TIM first needs a reliable and robust model to predict the required duration for clearance operations and assess its time-varying impacts on the conditions, because such information is essential for determining the proper control strategies and the sequences of the response tasks. Figure 6 shows those three key models and their interrelations in supporting the TIM's operations.

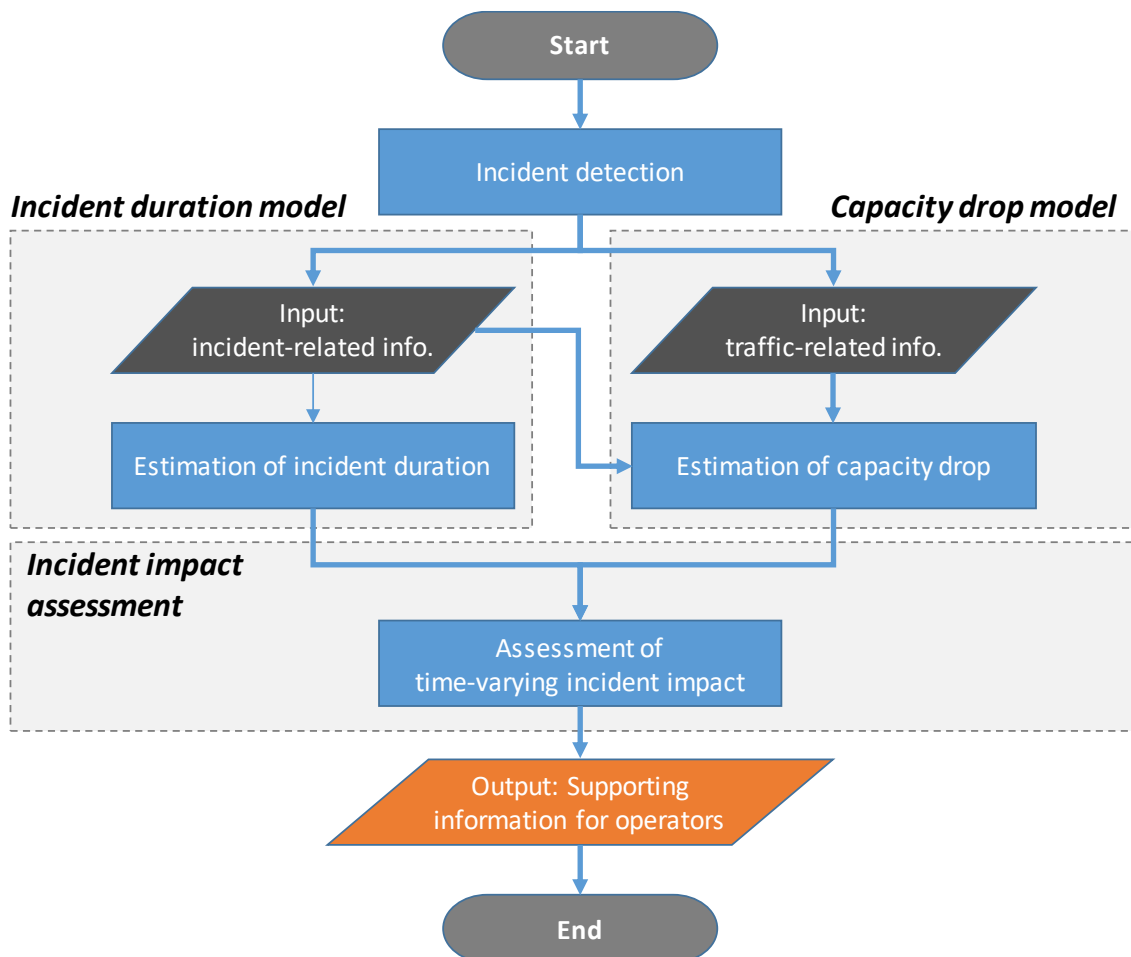


Figure 6. Framework of the proposed TIM support system

The first is the incident duration prediction model which is designed to estimate the duration of a detected incident using available information, such as incident type, the number of blocked lanes, the number of involved vehicles and response units, pavement conditions, time and location, etc. The second model is to estimate the reduced capacity of the roadway impacted by an incident, based on both the incident-related and the traffic-related information, such as traffic speed, volume, merging ratio, etc. The last model functions to predict its time-varying impacts, based on the estimates from the incident duration model and the capacity-reduction model. With all above models, the TIM support system can provide the temporal and spatial impacts of a detected incident to the control centers, allowing their operators/respondents to select proper operational strategies and execute all essential tasks in proper sequence.

3.3.1. Incident duration prediction models

To contend with different levels of data availability at different stages of incident response and clearance, this study proposes a multi-functional incident duration model to support the decision making of the operators in real time and detect faults and errors in the system. The proposed model, depending on the available information from the real-time operations, will provide the following three types of information with respect to a detected incident: interval estimate, point estimate, and outlier score. Figure 7 shows its core modeling concept and key system components.

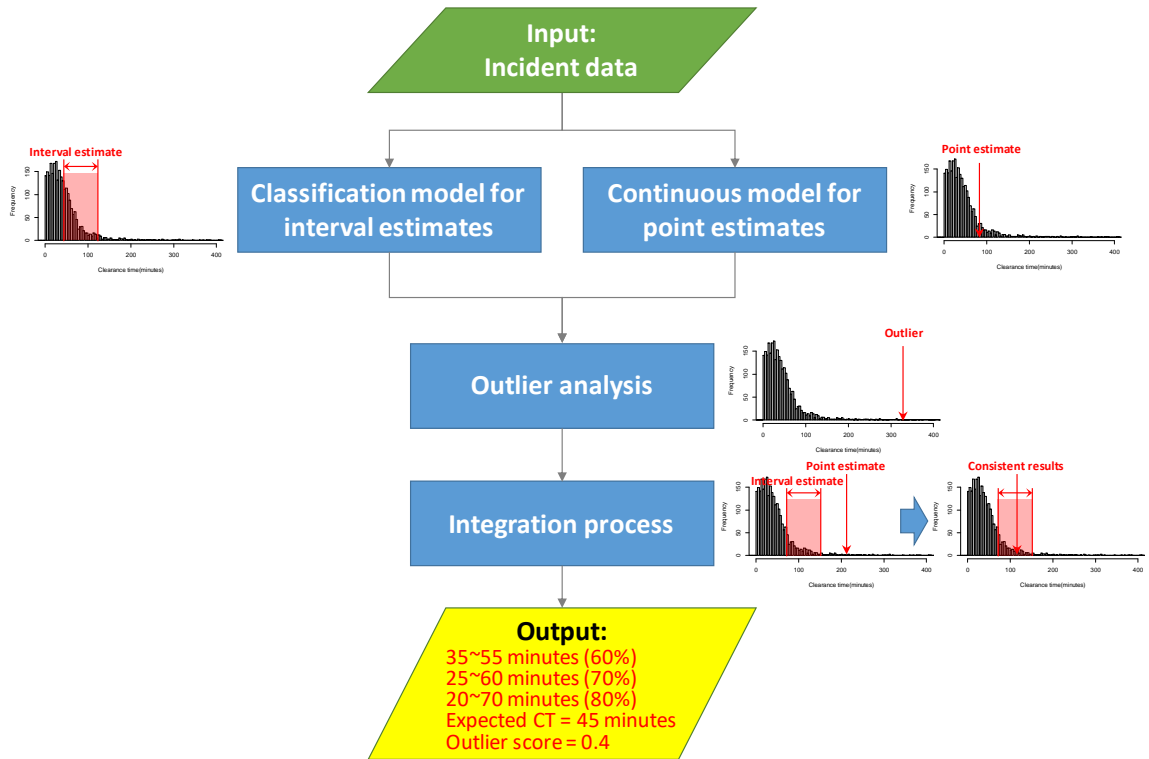


Figure 7. Components of the proposed incident duration prediction system

The first main component of the proposed incident duration prediction system is a classification model which functions to provide the estimated intervals with different confidence levels, using real-time available incident information. Its second component, a continuous model, is to provide the point-estimated duration of a detected incident, allowing users to assess the incident impact, and thus to determine the activation/deactivation time of all related control strategies. The last components is a set of supplemental modules to improve the estimated results from both the point and interval based prediction results and examine system deficiencies, based on some insightful information obtained from outlier analysis. Figure 8 shows how those components are designed to work together for the potential system users.

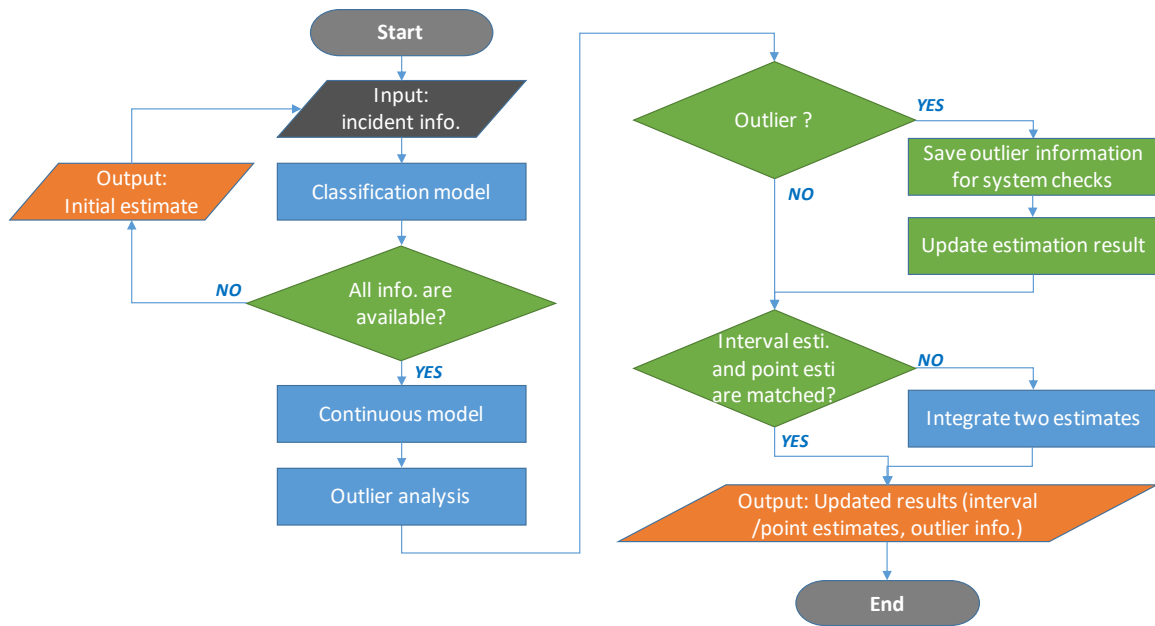


Figure 8. Framework of the proposed incident duration prediction system

In practice, the TMC system, based on the incident information from the scene, shall first apply the classification module in the prediction model to provide the estimated clearance intervals of a detected incident, and then proceed to a point-based estimate when more information becomes available. The outlier analysis process is used to evaluate if the detected incident, based on its nature and available information, may fall into the category of anomaly defined as outliers. If the incident is detected as an outlier, one shall update the interval and point estimates with additional rules developed from the prediction results from outlier analysis and forward that outlier information to the TMC for further analysis of system faults later. Lastly, when the interval and point estimates for the detected incident case are not consistent, the prediction system will apply the weighted logic to generate the integrated estimates from two estimates.

3.3.2. Capacity drop prediction and incident impact analysis models

To assess the impacts of a detected incident on traffic conditions such as the travel time delay and queues, the control center needs a capacity drop model to perform the analysis. As mentioned in the literature review, various mathematical and statistical capacity models have been developed by the traffic community. This study is focused on developing an analytical method for an efficient and convenient estimate of the capacity reduction and its resulting impacts on the traffic flows by a detected incident. Understanding key factors and their interrelations contributing to the capacity drop will allow the control center operators to better assess the spatial and temporal impacts of the detected incidents, and then to implement the most cost-effective control strategy in real time. Figure 9 shows the operational relations between its key components and system output.

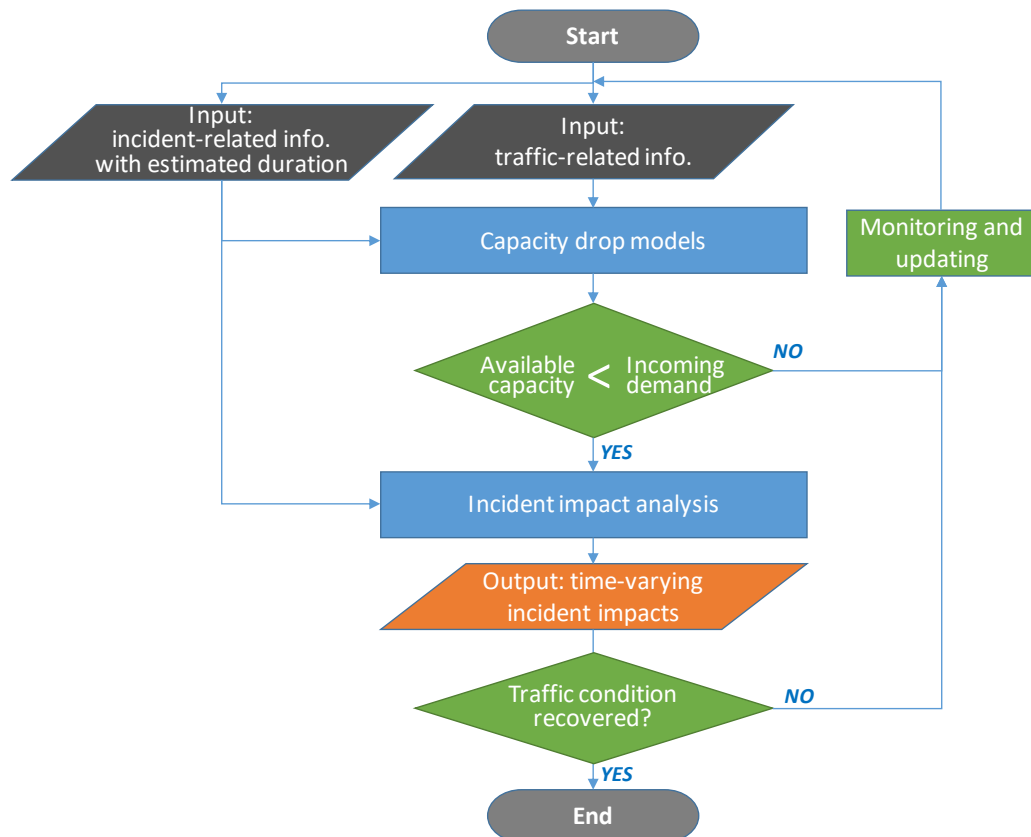


Figure 9. Framework of the proposed model for estimating the capacity reduction due to traffic incidents

Notably, its first task is to collect incident-related information and the estimated incident duration from the incident duration model, and then to receive the traffic-related data from the scene, including incoming traffic volume, vehicle speed, merging ratio, etc. With all such essential information, one can then execute the capacity reduction model to estimate the reduced capacity due to the incident, and then provide the predicted time-varying queue and delay information for the control center to decide when and where different control strategies ought to be deployed.

Chapter 4. Interval-based model for incident duration

4.1. Introduction

In review of the incident response practices in most state highway agencies [3], one may recognize that a variety of factors and their complex interrelations have limited the field applications of various existing models or tools for estimating incident duration. Examples of those challenges include: 1) a large number of variables (e.g., incident type, lane blockages, available resources) may have significant impacts on the clearance duration of a detected incident; 2) many of those critical factors, such as incident type (i.e., collision with fatality, collision with property damage, or collision with personal injury), responders (i.e., needs of tow units, medical units, and fireboard units), and environmental conditions, are qualitative in nature and often neither available nor collectable during the incident response and clearance process; 3) available models or tools may not be sufficiently robust to contend with the inevitable missing data or data errors due to human-factor related issues; 4) the quantitative output from some elegant advanced models may not directly correspond to the available management actions during incident response operations; and 5) most commonly and critically, many responsible highway agencies do not have an effective system to collect various critical data associated with incident response operations in real time at the desirable level of accuracy.

Therefore, the purpose of this chapter is developing a robust and practical incident duration model to provide a reliable and sufficient estimation information for the incident duration to the operators.

4.2. Data description and preprocessing

To develop an incident duration model, this study used the incident dataset from 2012–2017 in CHART II Database. The data from the first four years (2012–2015) were used for calibration, the data from 2016 were used for validation, and the data from 2017 were used for evaluation. As a prototype system, the spatial scope covers only MD I-95 from exit 27 to exit 109, and the data include only incidents with collision and travel lane blockage. This study defines the incident clearance time (CT) as the time elapsed from the arrival of the first response unit at the scene to the complete clearance of the incident. Note that the clearance duration is selected instead of the entire incident duration because most incident response times under the Maryland-CHART highway patrolling strategy are notably stable and only vary in a notably indifferent range. Table 1 shows the variables recorded during each incident event, which are the incident type, location, time, lane blockage information, involved vehicle information, environmental conditions, response unit information, etc.

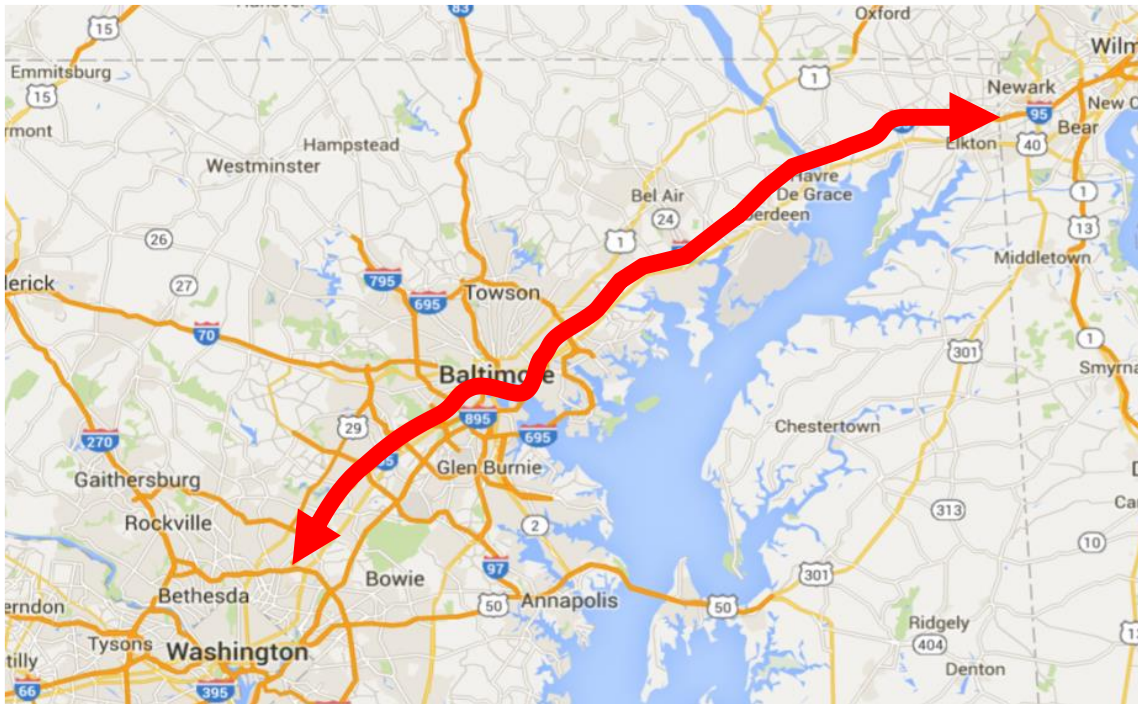


Figure 10. Study scope

Table 1. List of key factors associated with the response/clearance of a detected incident

Category	Variables	Classification
Incident type	Incident type	Collision with Fatality (CF), Collision with Personal Injury (CPI), Collision with Property Damage (CPD)
Time	Hour indicator	AM-peak (7AM~10AM), Day time (10AM~4PM), PM-peak (4PM~7PM), Night time (7PM~7AM)
	Weekend indicator	Weekend, Weekday
	Holiday indicator	Holiday, Non-Holiday
	Season indicator	Spring, Summer, Fall, Winter
Location	County indicator	Prince George, Howard, Baltimore, Baltimore City, Harford, Cecil
	Direction indicator	Northbound, Southbound
	Exit number indicator	Exit 27, ..., Exit 109
Environmental condition	Pavement condition indicator	Dry, Wet, Snow/Ice, Chemical wet, Unspecified
	Hazard material related	Yes, No
Operation center	Center indicator	AOC, TOC3, TOC4, TOC5, SOC, Others
Lane blockage	# of blocked lanes*	1, 2, 3, 4...
	# of blocked shoulders	0, 1, 2, 3...
	# of blocked travel lanes**	0, 1, 2, 3...
	# of blocked traffic lanes	0, 1, 2, 3...
	# of blocked auxiliary lanes***	0, 1, 2, 3...
	Travel lane blocked in tunnel	Yes, No
	Travel lane blocked in toll	Yes, No
Involved vehicle	Vehicle states	Jack-knifed, Over-turned, Lost-load
	# of total involved vehicles****	1, 2, 3, 4...
	# of involved passenger cars	0, 1, 2, 3...
	# of involved trucks	0, 1, 2, 3...
	# of involved motorcycles	0, 1, 2, 3...
Responder	# of total response units	1, 2, 3, 4...
	# of arrived CHART	0, 1, 2, 3...
	# of arrived Police	0, 1, 2, 3...
	# of arrived Fireboard	0, 1, 2, 3...
	# of arrived Medical service	0, 1, 2, 3...
	# of arrived TOW service	0, 1, 2, 3...
	FIRST responder	CHART, Police, Fireboard, Medical, Tow

* Lanes = Shoulders + Travel lanes

** Travel lanes = Traffic lanes + Auxiliary lanes

*** Auxiliary lane includes on-ramp, off-ramp, acceleration lane, deceleration lane, and collector/distributor lane

**** Vehicle includes passenger car, truck, bus, cyclist, pedestrian, and motorcycle

Table 2. Sample sizes by lane blockage status and years (Collision)

Year	2012	2013	2014	2015	2016	2017	total
Travel lane blockage	328 (33.3%)	437 (37.4%)	595 (34.2%)	598 (31.2%)	626 (31.5%)	444 (24.9%)	3028 (31.6%)
<i>NO Travel lane blockage</i>	377 (38.2%)	364 (31.1%)	611 (35.1%)	806 (42.1%)	800 (40.3%)	693 (38.9%)	3651 (38.1%)
<i>NO lane blockage information</i>	281 (28.5%)	368 (31.5%)	536 (30.8%)	512 (26.7%)	560 (28.2%)	644 (36.2%)	2901 (30.3%)
Total (percentage)	986 (100%)	1169 (100%)	1742 (100%)	1916 (100%)	1986 (100%)	1781 (100%)	9580 (100%)

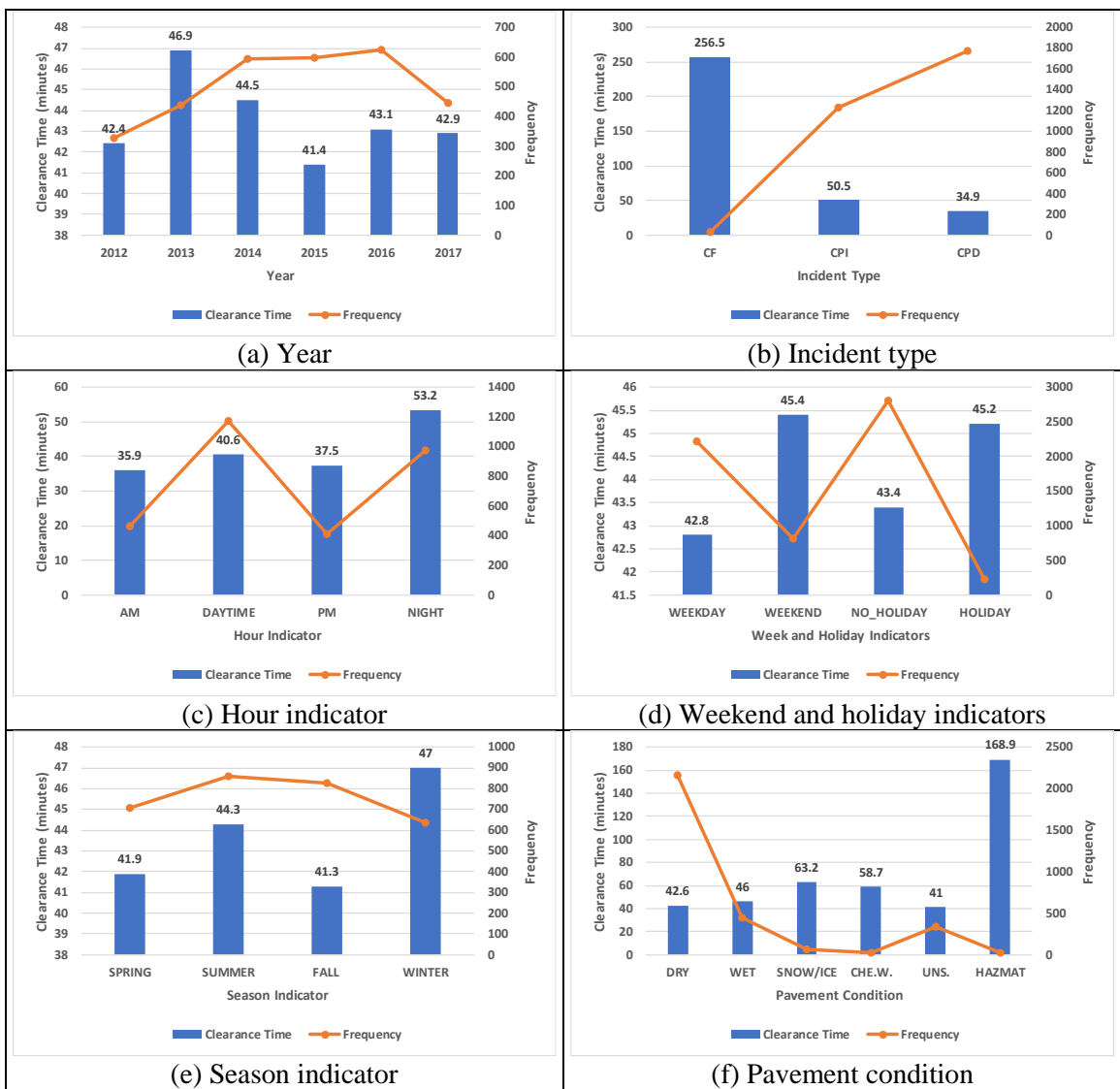


Figure 11. Incident clearance time and frequency by year, incident type, hour, weekend, holiday, season, and pavement condition (Collision with travel lane blockage)

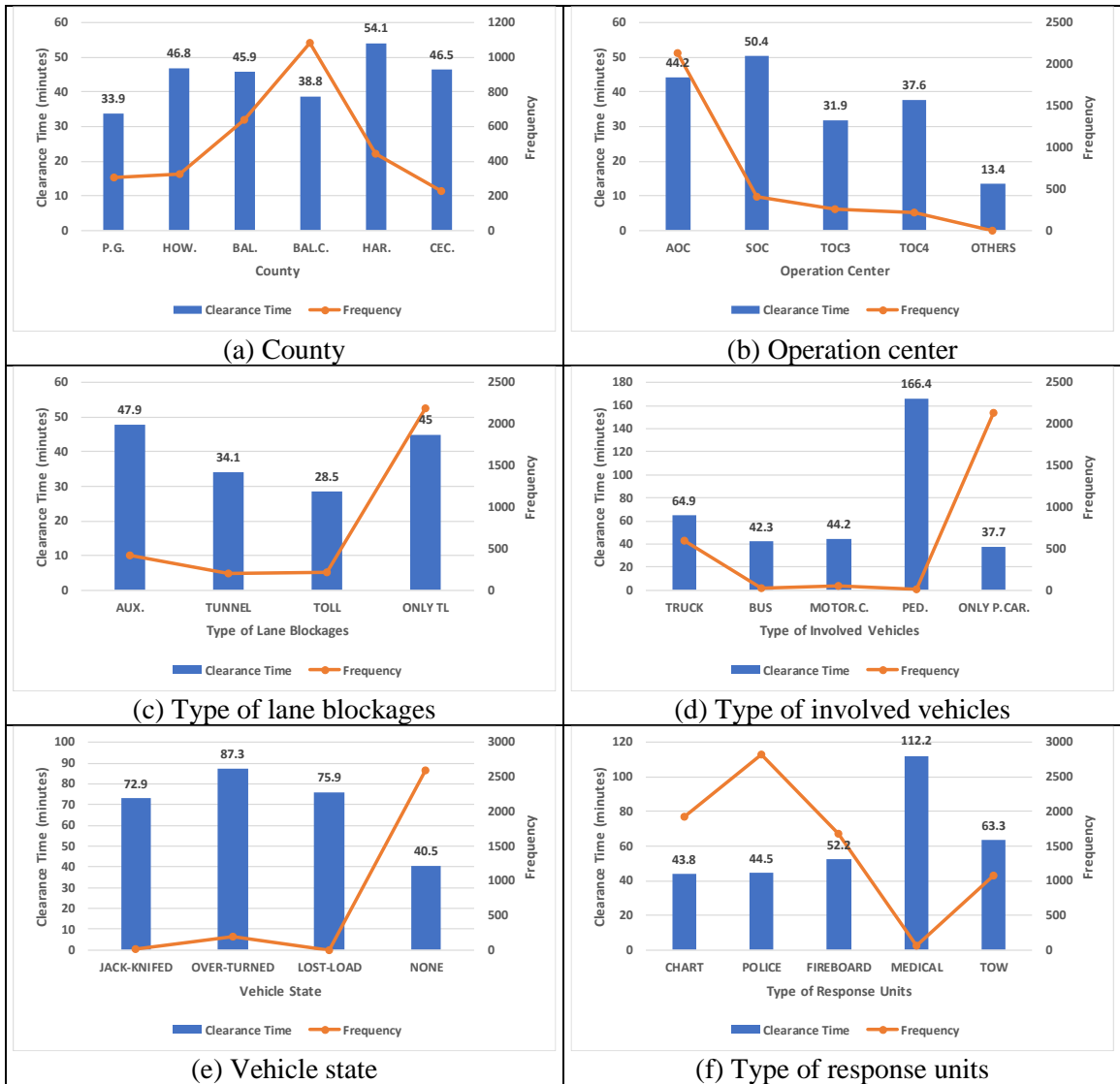


Figure 12. Incident clearance time and frequency by county, operation center, lane blockage, involved vehicle, vehicle state, and response units (Collision with travel lane blockage)

Table 2, Figure 11, and Figure 12 show the characteristics of the incident data in this study area. As shown in Table 2, among the incident datasets in the years 2012-2017, around 30 percent of the collision cases have travel-lane blockages, and the other cases have no travel-lane blockages or no information for lane status. The details for incident duration and frequency of the collision-with-travel-lane-blockage cases are as follows:

- **Mean of incident clearance time and frequency** (Figure 11 (a)): the mean of incident clearance time of each year ranges from 41.4 to 46.9 minutes, i.e., less than 6 minutes difference in each year. The number of cases is also similar in each year.
- **Incident type** (Figure 11 (b)): the fatality-related incidents show the longest clearance time as about four hours and a half, despite its very small portion in the dataset. The injury-related incidents usually take more fifteen minutes than the no-injury-related incidents.
- **Hour indicator** (Figure 11 (c)): incidents in the nighttime take around more 13 minutes than in the daytime. Interestingly, the incidents in the peak-hours are cleared quicker than the other incidents in the daytime. It seems that highway agencies operate more response resources and further effort to clear them during those peak-hours, because of its expected larger impacts on the networks [57].
- **Weekend and holiday indicators** (Figure 11 (d)): incidents during the weekend and holiday usually take longer clearance time than do in the other days, because of unusual heavy traffics or the lack of operational resources such as TOW services in those periods [10].
- **Season indicator** (Figure 11 (e)): incidents in the winter season have a longer clearance time than in the other seasons. It seems to be a bad road conditions such as snowing or icing conditions [10], or the sudden increase of accidents in the heavy snowing days and its resulting operating issues (e.g., December 8, Sunday, 2013; January 20, Wednesday, 2016 in the study area).
- **Pavement condition** (Figure 11 (f)): the hazard-material-related incidents take around three hours to clear them, due to their required special treatments. Also,

incidents under the snow/ice and chemical-wet (e.g., fuel leak) conditions take longer than under the other conditions.

- **County** (Figure 12 (a)): incidents in the different locations show the different incident clearance time. Incidents in Harford county show the longest clearance time and incidents in Prince George's county show the shortest clearance time.
- **Operation center** (Figure 12 (b)): incidents involved in the different operation centers also show the different incident clearance time. It seems to be caused by the different performances or control areas of each operation center.
- **Type of lane blockages** (Figure 12 (c)): incidents with the auxiliary-lane-blockage usually take a little bit longer than the other incidents with only-travel-lane-blockage. Incidents with the toll-lane or tunnel blockage cases take shorter than the other incidents with only-travel-lane-blockage.
- **Type of involved vehicles** (Figure 12 (d)): usually, truck-involved incidents take a longer clearance time than the only-passenger-car-involved incidents. Interestingly, when a pedestrian is involved in the incident, it takes more than two hours to clear. Because most pedestrian-involved incidents are fatality-accidents and they seem to need an additional investigation.
- **Vehicle status** (Figure 12 (e)): when an accident vehicle is jack-knifed, over-turned, or lost-load, it takes about more twice than does the normal incident without those special situations.
- **Type of response units** (Figure 12 (f)): when a medical service is required, the incident takes the longest clearance time since its severity. Also, the incident required the TOW service takes a little bit longer clearance time, because the time of TOW

services to arrive the scene usually is longer than that of the other response units, and they need an additional time to move the accident vehicle using their equipment [10].

To remove potential data errors by system operators, often occurred during real-time incident response and management process, this study has taken the following steps after consulting with field operators. But, note that this study has adjusted the training dataset to remove the potential data errors because there is no way to know the ground-truth value during prediction process, so this data preprocessing is not available during the prediction process.

- **Step-1:** If the difference between the event-cleared time and the all-blocked-lane-reopened time, including shoulders, is larger than 5 minutes (see Figure 13 (a)), one takes the all-blocked-lane-reopened time as the event-cleared time in computing the incident clearance time.
- **Step-2:** If the difference between the all-blocked-lane-reopened time and the all-blocked-travel-lane-reopened time is larger than 32 minutes, one uses the sum of the all-travel-lane-reopened time plus 32 minutes as the event-cleared time in computing the incident clearance time. The logic behind this step can be seen from Figure 13 (b) where the 90% of the open time of all blocked lanes, including shoulders, for all recorded incidents are within 32 minutes difference from the open time of blocked travel lanes.
- **Step-3:** Since some recorded incident events show an unreasonably short clearance time due to various reasons, one uses the 10% quantile as the lower bound to trim the available data.

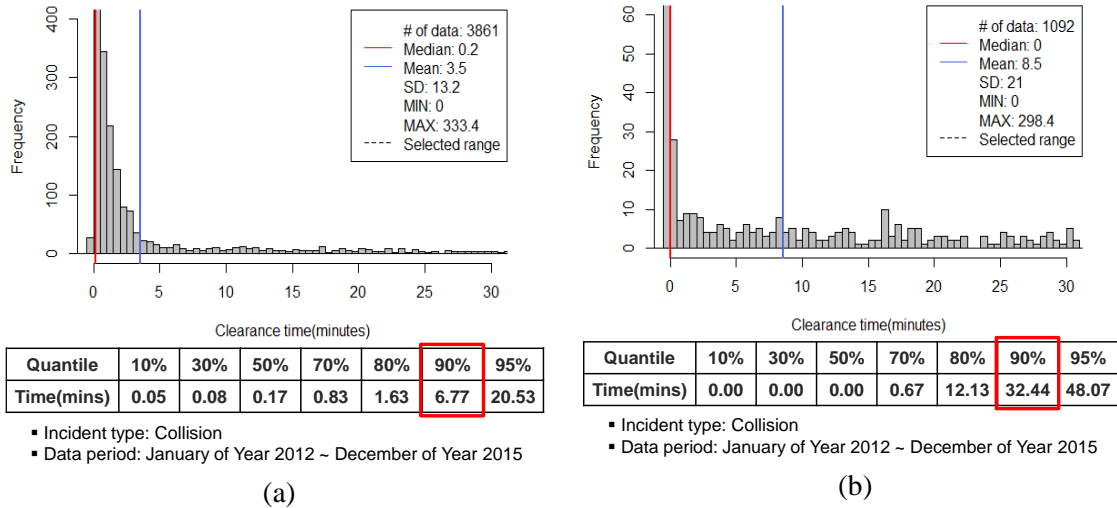


Figure 13. (a) Distribution of time differences between the all-lane-reopen times and the event-cleared times, (b) distribution of time differences between the travel-lane-reopen times and the all-lane-reopen times.

The final training dataset after the data preprocessing is shown in Figure 14.

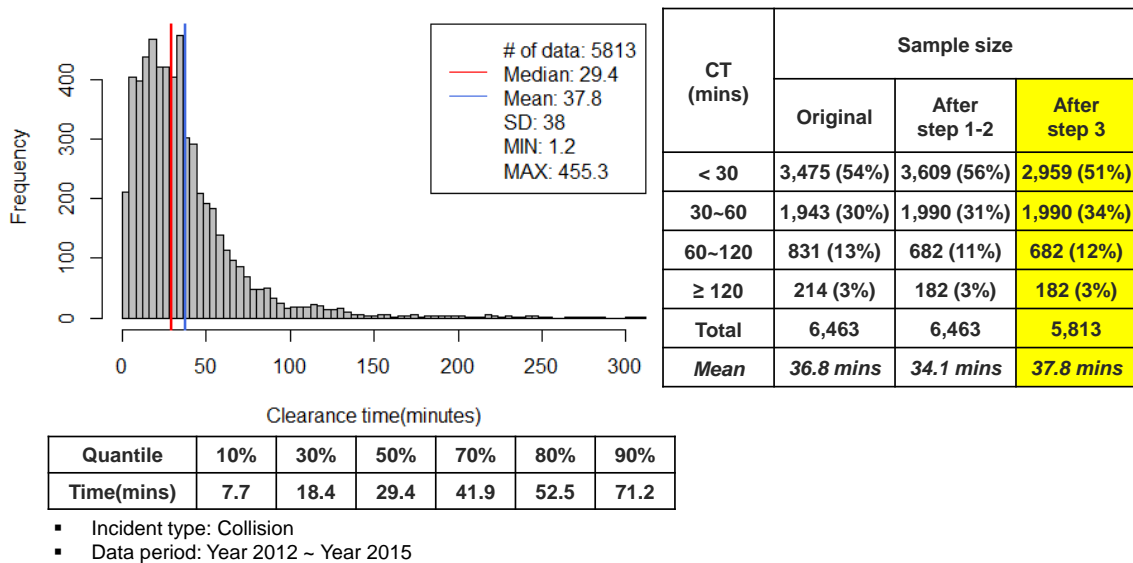


Figure 14. Distribution of clearance times after the data preprocessing

4.3. Modeling concept and methodology

The incident duration data have a right-skewed distribution, many qualitative variables, and outliers [10]. Many studies have provided reliable estimation results to predict incident duration. However, in the operational viewpoint, it is still limited to apply these incident duration models to the field even when a model with reliable performance is developed. One of the reasons is that some input values of the model to estimate the incident clearance time are not always available in the field, but most models in previous studies require all input information to estimate the incident clearance time. Thus, these models cannot provide the initial incident duration information to the operators based on incomplete incoming information from the scene. The operators should wait for all input information from the field to estimate an expected incident duration and deploy operational strategies according to the incident duration provided by the prediction model. For immediate responses, this issue is critical.

Second, the operators cannot anticipate the risk for the prediction error of the prediction model. Some models provide an exact number for expected incident duration. Other models provide a range of the expected incident duration. Regardless, most incident duration models provide one piece of information of the expected incident duration based on their estimation results, which does not include any risk information for misprediction. If a model can provide an expected incident duration or different time ranges with probabilities for prediction error, the operators can anticipate the risk of misprediction. These information and informative methods will be useful for the operators to make a decision for various operational strategies.

Thus, the purpose of this section is to develop an incident duration model with high field applicability using a hybrid method, which provides different incident duration ranges with different confidence levels and reliable estimation information even though all information from the scene is not available.

4.4. Model development and estimation results

In view of the large number of qualitative factors, this study first categorized the incident data based on those major factors, and then searched for classification rules to divide the incident data into several categories based on incident clearance time (CT), using the association rule mining method and hybrid modeling method which are well fit to the knowledge-based modeling. Association rule mining method is commonly used to analyze data for frequent if/then patterns and to identify the most important relationships [114, 115]. It has the advantages for users to interpret the resulting relationships and to implement the results in practice. Hybrid modeling method is mainly selected to integrate different methods based on incident duration natures [3]. It enables the developed system to provide the estimate in a sequential manner, and to deal with an incident with missing data. Moreover, from each category of incident clearance time, this study has selected several ranges of estimated clearance duration according to approximately 60%, 70%, and 80% of probability for field operators to make a proper assessment. Note that the selection of confidence thresholds can be specified by the users based on the available incident response resources and the quality of available data recorded during the incident clearance process.

4.4.1. Incident categorization

Note that key factors associated with the resulting incident duration have complex relationships with one another. To sort out such relationships, all available incident data are categorized with the following steps: 1) dividing all incident data into two categories: shoulder-only-blockage and travel-lane-blockage; 2) providing the mean incident clearance time and its estimated ranges with approximately 60%, 70%, and 80% of probability; and 3) further dividing the travel-lane-blockage cases into several groups, based on the incident types and number of blocked travel lanes. Note that no further classification for shoulder-only-blockage cases was taken, because they all lie in a very stable and small range. By the same token, the incident data for Collision with Fatality (CF), resulted in the travel-lane-blockage, was not further classified with the number of blocked lanes due to the small sample size.

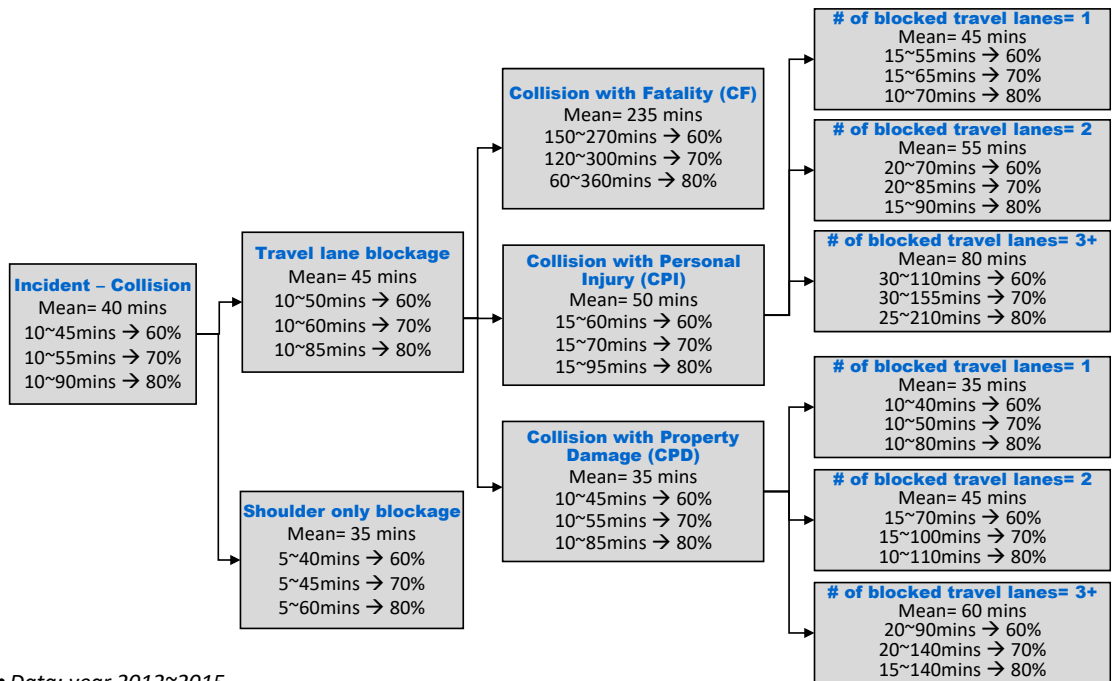


Figure 15. Initial incident categorization and estimated clearance duration

Such an initial estimate for each detected incident (see Figure 15) shall be useful for the operators to assess the potential impacts when the information for the full picture of the detected incident is not available in the clearance process yet.

4.4.2. Classification rules mining

Based on the incident categorization in Figure 15, one can provide the initial incident duration information for each incident case. In the subsequent steps, both the association mining method and hybrid modeling method are applied to the last seven categories in the travel-lane-blockage cases to increase the estimation accuracy. Note that the clearance duration for most shoulder-only blockage cases is around the interval of 30 minutes, so this study only considers the travel-lane-blockage cases in exercising these modeling steps for the further classification.

4.4.2.1. Collision with Personal Injury (CPI) and Collision with Property Damage (CPD) in the Travel Lane Blockage Cases

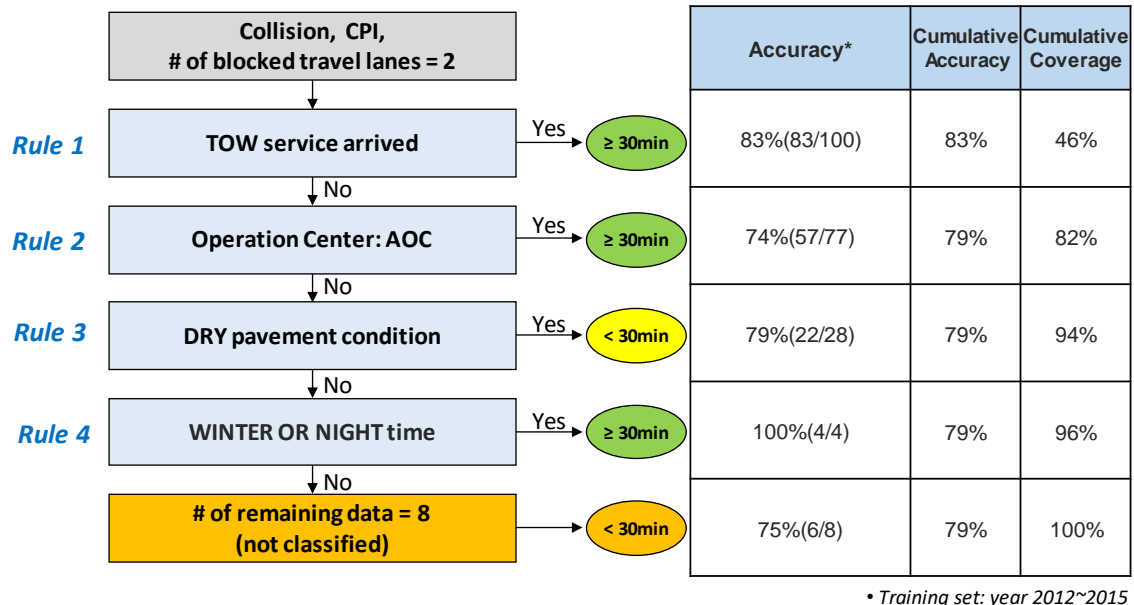
There are six categories of CPI and CPD in the travel-lane-blockage cases, which are divided by the number of blocked travel lanes. To estimate the range of incident clearance time for each category for operation's needs, this study first classifies the data into "< 30 minutes" and "≥ 30 minutes" using the association rule mining method. Then, from the "≥ 30 minutes" subset, one searches other classification rules to classify the data into "< 60 minutes" and "≥ 60 minutes". Similarly, one can find classification rules to further categorize the expected incident clearance times for "< 120 minutes" and "≥ 120 minutes".

Finally, based on the distributions of these categories, one produces the three intervals of expected incident clearance time under the 60%, 70%, and 80% confidence levels.

The procedures to model the CPI and CPD in the travel-lane-blockage events to generate classification rules for “< 30 minutes” and “≥ 30 minutes” as follows:

- **Step-1:** From the dataset, search for classification rules to classify the data into “< 30 minutes” and “≥ 30 minutes” using the association rule mining method.
- **Step-2:** Select a critical rule with approximately more than 75% confidence level and the highest support level.
- **Step-3:** Filter out the data associated with the selected rule from the dataset and proceed with further classification for the remaining data.
- **Step-4:** Stop the classification procedures if no further rule for classification can be derived from the remaining data.
- **Step-5:** Otherwise, go to Step-1 and repeat the same procedures.

Figure 16 shows an example of CPI with two blocked travel lanes. Using such hybrid rules for classification, one finds four classification rules for “< 30 minutes” and “≥ 30 minutes” and set up the estimated interval for incident clearance time for the remaining data.



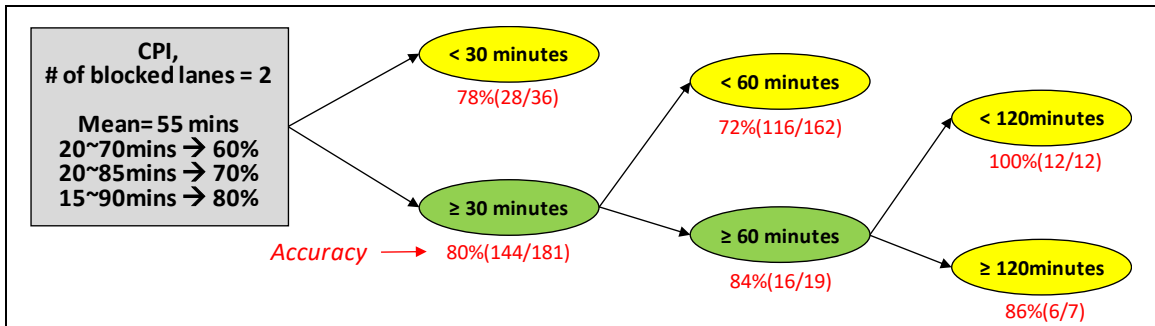
* Numbers in parenthesis represent “the number of data that are correctly estimated by the rule in the remaining dataset / the number of data that is categorized by the rule in the remaining dataset”

Figure 16. An example of the classification rules for “< 30 mins” and “≥ 30 mins” in the CPI with two blocked travel lanes

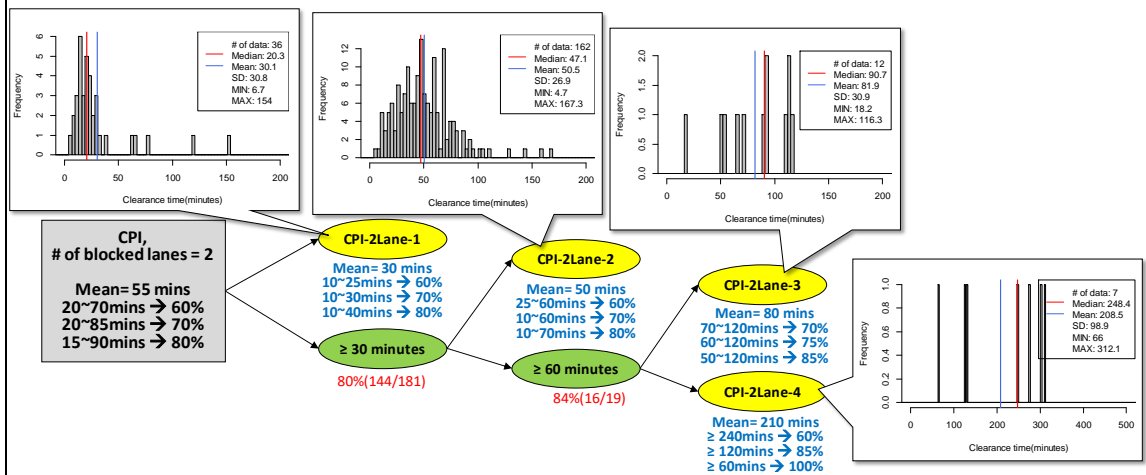
In the next step, this study applied the same classification procedures to the dataset classified in the subset of “≥ 30 minutes” in the previous step to further divide the data into “< 60 minutes” and “≥ 60 minutes”. Same procedures were also employed for “< 120 minutes” and “≥ 120 minutes” categories. Figure 17 shows the results of this hybrid classification procedure for six intervals of estimated incident clearance time, such as “< 30 minutes”, “≥ 30 minutes”, “< 60 minutes”, “≥ 60 minutes”, “< 120 minutes” and “≥ 120 minutes”, as an example of application for CPI with two-travel-lane blockage.

Figure 17 (c) shows the rule application process for the CPI with two blocked travel lanes. If an incident with CPI and two blocked travel lanes was reported, this event will go to the rules in the first table in Figure 17 (c). Based on the rules in the first table, this event will be estimated to belong to “CPI-2Lane-1” or “CPI-2Lane-2”. If this event is confirmed

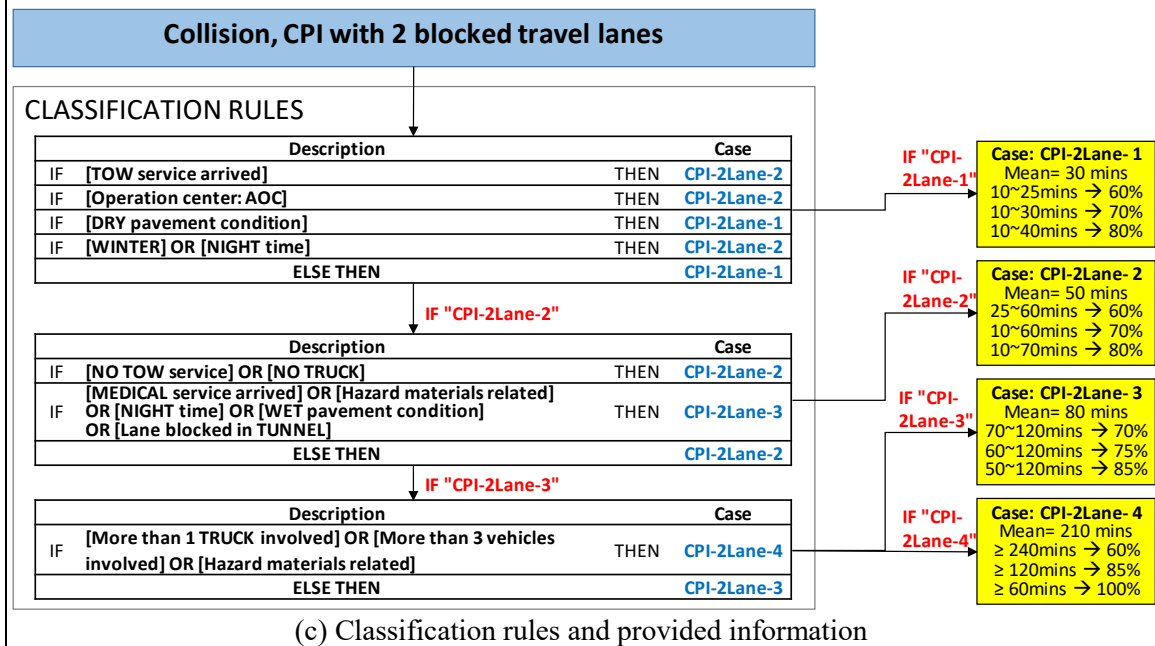
from the field report to belong to “CPI-2Lane-1”, the model will then provide the expected incident clearance time information associated with “CPI-2Lane-1”. If not, this event will go to the rules in the second table for further classification. This is the process to provide the expected incident clearance time information sequentially using the proposed system. Note that this hybrid classification model is to be executed on a sequential process, so it can provide an initial estimate in real time, even if full information for a detected incident from the field is not available.



(a) Six expected incident clearance time ranges



(b) Distributions and ranges of incident clearance time with different confidence levels



(c) Classification rules and provided information

Figure 17. An example of application for CPI with two-travel-lane blockage

4.4.2.2. Collision with Fatality (CF) in the Travel Lane Blockage Cases

There are only 25 in the training set (2012~2015) for this type of incidents. Thus, the modeling procedure for the CPI and CPD in the travel lane blockage cases is not suitable for those CF cases. In addition, most CF cases have incident clearance times longer than 120 minutes, so different boundaries were set for the clearance time ranges for such incidents. Rule-generation procedures used for analyzing CF incidents are summarized below:

- **Step-1:** Calculate the median value of the incident clearance time of the dataset and divide the dataset into two groups using the median value.
- **Step-2:** Search the classification rules to classify the data into those two groups using the association rule mining method.
- **Step-3:** Select a critical rule with approximately more than the 70% confidence level and the highest support level.
- **Step-4:** Divide the data into two subsets according to the selected rules and set up the estimated incident clearance time ranges, based on the distribution of incidents in each subset.
- **Step-5:** Stop the procedures if the estimated incident clearance time range is sufficiently robust for use by control center operators.
- **Step-6:** Otherwise, go to the first step and repeat the procedures for incident data in each subset.

This classification mining method can sequentially classify the dataset into two groups, based on the median value. Thus, it can continuously reduce the distribution and

range of the incident clearance times from each subset. Figure 18 shows the classification results and the estimated intervals for incident clearance time. Most estimated ranges for clearance time for collision-fatality incidents have an interval of 60 minutes with approximately 70% of confidence.

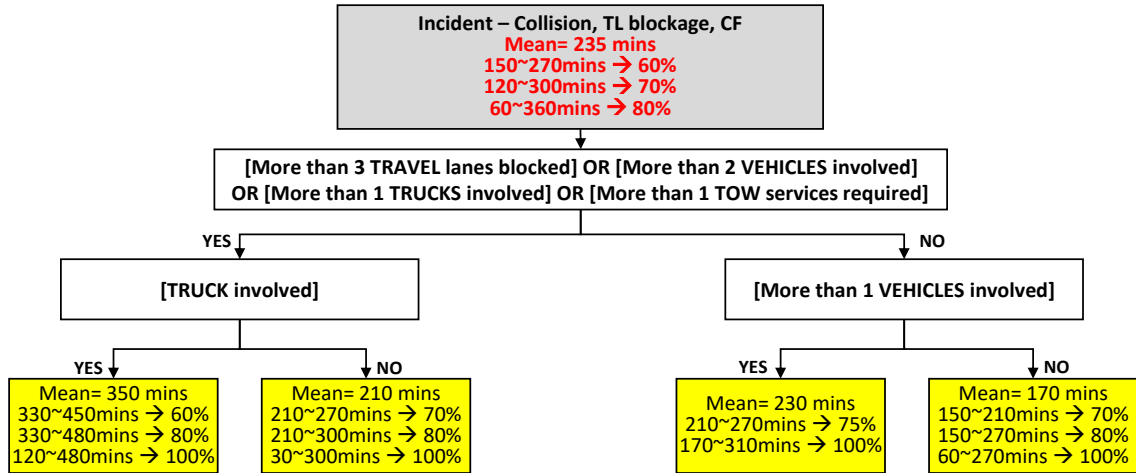


Figure 18. Classification rules and incident clearance time ranges in the CF with travel lane blockage

4.5. Model evaluation and discussion

This study has used the data over four years (2012~2015) as the training set for model calibration. Based on the developed model results, this study also has evaluated the model using the data in year 2016 and 2017 as the test set. Table 3 shows the evaluation result by incident types and the number of blocked travel lanes, and Table 4 shows its performance by different categories of the actual incident clearance time. Lastly Table 5 shows its distribution of estimation errors.

Table 3. Accuracy of interval estimates by incident types and # of blocked travel lanes

	Travel lane (TL) blockage cases							
	CF	CPI			CPD			Total
		1 TL block.	2	3+	1	2	3+	
Training set (2012~2015)	100% (21/21)	77.7% (363/467)	83.6% (173/207)	78.7% (70/89)	73.9% (645/873)	80.7% (134/166)	81.1% (30/37)	77.2% (1436/1860)
Test set (2017)	75.0% (3/4)	74.8% (83/111)	90.9% (30/33)	80.0% (12/15)	76.1% (150/197)	79.6% (43/54)	91.7% (11/12)	77.9% (332/426)

* Numbers in parenthesis represent “the number of data which is correctly estimated by the model / the total number of data in the category”

Table 4. Accuracy of interval estimates by CT categories

Dataset	Actual CT (minutes)				Overall
	< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	96.1% (719/748)	81.6% (549/673)	35.7% (122/342)	47.4% (46/97)	77.2% (1436/1860)
Test set (2017)	97.0% (161/166)	83.2% (129/155)	33.8% (27/80)	60.0% (15/25)	77.9% (332/426)

Table 5. Distribution of interval estimation errors by CT categories

Errors		Actual CT (minutes)								Overall	
		< 30		30 ~ 60		60 ~ 120		≥ 120			
		TRAIN	TEST	TRAIN	TEST	TRAIN	TEST	TRAIN	TEST	TRAIN	TEST
Overestimated	> 120	0	0	0	0	0	0	0	0	0	0
	60~120	0	0	0	0	0	0	0	0	0	0
	30~60	0	0	0	0	1	0	0	0	1	0
	10~30	0	0	2	1	0	0	0	0	2	1
	0~10	29	5	0	2	1	1	0	0	30	8
Within boundaries		719	161	549	129	122	27	46	15	1436	332
Underestimated	-10~0	0	0	76	11	59	19	2	2	137	32
	-30~-10	0	0	46	12	86	17	3	0	135	29
	-60~-30	0	0	0	0	60	13	5	3	65	16
	-120~-60	0	0	0	0	13	3	32	4	45	7
	< -120	0	0	0	0	0	0	9	1	9	1
Total # of cases		748	166	673	155	342	80	97	25	1860	426
TP rate and Accuracy		96.1%	97.0%	81.6%	83.2%	35.7%	33.8%	47.4%	60.0%	77.2%	77.9%

* TRAIN: training set (2012-2015) / TEST: test set (2017)

As seen in Table 3, the initial classification model can obtain more than 75% of accuracy in the training and test sets, and those results are quite consistent across the different incident types and travel lane blockage status. In Table 4 and Table 5, the category

for 60~120 minutes shows the lowest accuracy, because the incident data have a highly right-skewed distribution with a long tail on the right-side. However, more than 96% (406/422) of the cases with 60~120 minutes have less than 60 minutes time difference between the actual clearance time and estimated clearance time in the training set and test set, and the overall accuracies are also quite robust to provide a reliable prediction information.

The proposed model features its use of interval-based estimates, derived from the knowledge of the historical incident response data with different confidence levels for incident clearance time, and the rule-based structure for providing an initial estimate of the incident duration when all incident-related information is not available in the field. Its sequential nature allows the users to dynamically revise the estimate when additional data have been reported, because some key variables, affecting the duration of a detected incident, often only become available as the clearance operations progress.

Our preliminary evaluation results have shown the promise of the proposed system which with its invaluable historical knowledge can circumvent many data quality and availability issues plaguing the applicability of some state-of-the-art models on this subject. The proposed classification model can also be used to estimate reliable travel time during incidents, given the approximated or observed queue length. Furthermore, our proposed interval-based estimates with three thresholds (i.e., 30, 60, and 120 minutes) offer the information consistent with the potential ATIS/ATMS strategies (e.g., providing incident information only, advisory or mandatory detouring operations) for the responsible highway agencies to interact with en-route motorists and general public.

Chapter 5. Point-based model for incident duration

5.1. Modeling concept and methodology

Despite of the promotion of an interval-based estimate for incident duration, the information of a point estimate remains in high demand for many reasons in the fields. First, such information is convenient for computing the queue length and for determining a control point for various detour strategies. It is also notably useful and critical for travel time prediction and selection of strategies to dispatch the response units.

Many studies in the literature have developed the incident duration model to provide a reliable point-based estimate, which range from continuous models to hazard-based models [7, 17, 18, 21, 22, 27, 28, 31, 33], including quantile regression models [32, 34], rule-based models [42, 45] and hybrid models [10, 55-60, 62, 116]. However, despite their considerable efforts, most existing incident datasets share the following common issues that need to be addressed: 1) using one single statistical model is difficult to capture complex correlations and multicollinearity among some contributing factors; 2) there exists a tail risk due to the unique distribution of incident data; and 3) some critical factors, mostly associated with a long duration but a small sample size, are often dismissed, because most modeling algorithms focus on maximizing the overall accuracy.

Therefore, the purpose of this section is to develop a robust model for producing the point-based estimate for the duration of a detected incident.

5.2. Model development

As revealed in [10, 57], most recorded incident duration data exhibit a highly skewed distribution and have the following unique characteristics: 1) the resulting duration varies

significantly among incidents of different natures (e.g., collision only vs. involving fatality); 2) the required duration to clear each type of incidents is often dominated by one or two factors (e.g., truck over-turned or hazard material related); 3) some types of incidents, especially those in need of special equipment or assists, often have relative small samples in the database; and 4) some qualitative and quantitative factors recorded in the incident database are highly correlated or mutual dependent. In brief, in view of the unique distribution characteristics of incident data and the strengths of each available estimation method, one may effectively divide the entire dataset sequentially into several distinct subsets, allowing each to be best calibrated with the method, so as to yield the most robust estimate.

As stated previously, to develop a robust model for predicting a point-based incident duration for an available dataset, one needs to recognize and overcome the following vital issues: 1) incident nature (e.g., fatality) that dictates the resource needs and resulting duration for clearance operations should be taken into in selection of estimation models; 2) some variables may have quite different impacts on the resulting clearance duration of incidents with different natures; and 3) the estimation technique that can best fit each type of incidents may also vary with its available sample size.

With the aforementioned critical issues in mind, this study has proposed the following procedures to yield a reliable ensemble model for estimating incident duration:

- **Stage-1:** Partitioning of the entire dataset into several subsets via a sequential screening process, based on those factors contributing most to the duration of incidents with some unique natures.

- **Stage-2:** Applying different modeling methods to each subset, based on available sample sizes, distribution patterns, and key contributing factors.
- **Stage-3:** Integrating all those models, each developed specifically for one subset of incidents, with proper weights to constitute an ensemble model for the final estimate of a detected incident's required duration.

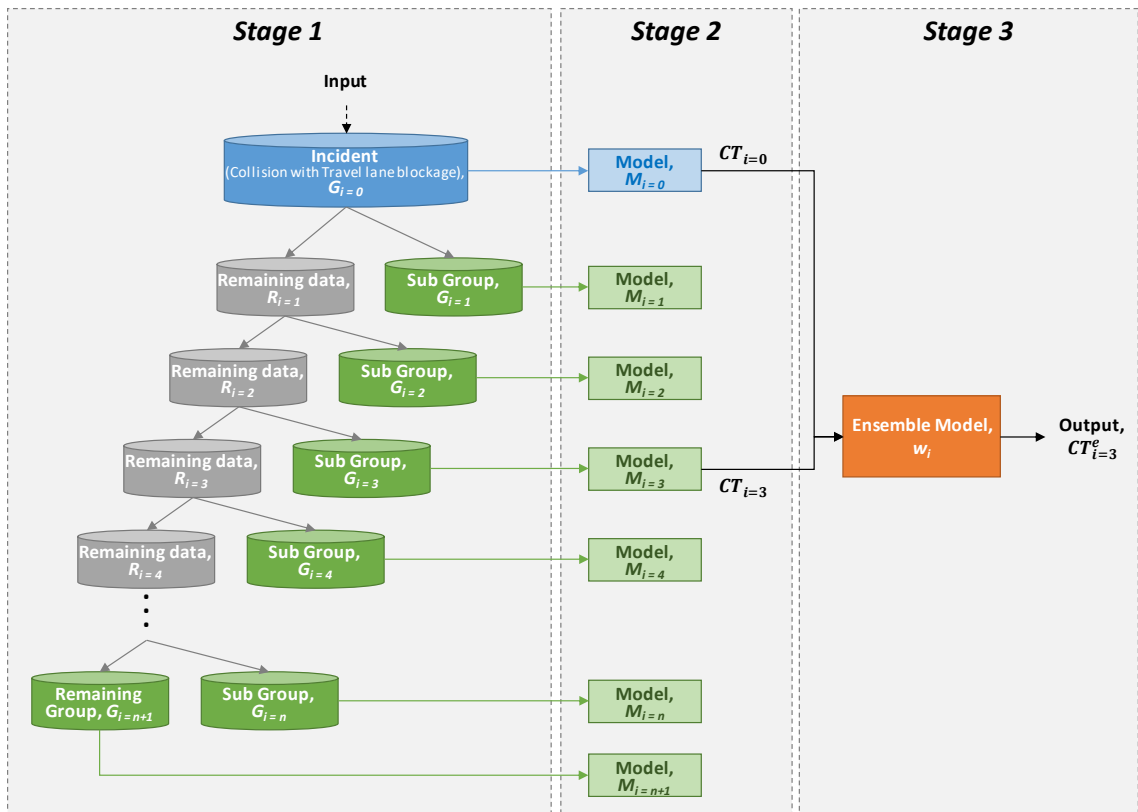


Figure 19. Graphically illustration of the modeling concept and its procedures

A graphical illustration of the interrelations between key activities conducted at each stage is shown in Figure 19, where i denotes a data group sequence, G_i is a data group in the sequence i , R_i is a remaining data group in the sequence i , M_i is an estimated incident duration model for each G_i , CT_i is an incident duration estimate using a model M_i , w_i is a weight for each estimate CT_i , and CT_i^e is a final ensemble estimate of the incident

belonging to a sequence data group i . A detailed discussion of key activities to be performed at each stage is presented in sequence below:

5.2.1. Stage-1: Sequential partitioning

The primary purpose at this stage is to divide the available dataset into several subsets, where most incidents' resulting clearance times in the same subset suffer mainly from the same common contributing factors. Several key steps to be done at this stage include:

- **Step-1:** Compute the median value for the entire dataset and for each subset that was classified with each of those identified key contributing factors. For instance, the median is 34 minutes for the overall samples of 2009 incidents, 236 minutes for the subset of 21 incidents with fatality, 73 minutes for incidents in need of tow services, and 100 minutes for those involving hazardous materials.
- **Step-2:** Divide the entire dataset into a subgroup (G_1) corresponding to the selected factor (f_1) that has the longest median value (e.g., the group of incidents with fatality in this dataset) and the remaining data (R_1) as show in Figure 19.
- **Step-3:** Follow the same logic to identify the most critical factor in the remaining dataset (R_1), and then re-divide the remaining dataset into a subgroup (G_2) and sub-dataset (R_2).
- **Step-4:** Combine the subgroup (G_i) with its previously identified subgroup (G_{i-1}) if their differences in median clearance duration are statistically insignificant (e.g., test with Mann–Whitney U test); and the data points in those two sequentially classified subgroups are highly correlated (e.g., with Spearman's correlation test).

- **Step-5:** Stop the procedures and assign the last remaining data ($R_{i=n}$) to the last remaining group ($G_{i=n+1}$), if no other contributing factor can be used for data classification, or no significant median differences exist between the subgroup (G_i) and the remaining data group (R_i).

Upon completing all activities in Stage-1, one can divide the entire dataset into several distinct subgroups with unique characteristics, and each can then be estimated with the most effective modeling method.

5.2.2. Stage-2: Modeling for each category

To reflect the significant impact of an identified factor (e.g., # of trucks) on certain types of incidents (e.g., incidents in need of tow services) and also to tackle the sample size as well as data distribution constraints, one can take the following steps in selecting and performing the estimation for incidents classified in each subgroup:

- **Step-1:** Use the rule-based method [10] to capture the relations between incident clearance time and its associated factors for those groups of a small sample size [117].
- **Step-2:** Directly apply the median value and its percentile interval for clearance time estimation for those groups suffering from both the small sample size and lack of definitive rules.
- **Step-3:** Perform the normality test with both the Kolmogorov-Smirnov and Shapiro-Wilk normality tests for those groups with a sufficient sample size.

Note that Figure 20 shows an example of the normality test, where the distribution and Q-Q plot of the incident clearance time for “TOW service arrived” group show a right-skewed shape. Since its P-values from the two normality tests are also smaller than the significant level ($\alpha=0.05$). Thus, one can conclude that the subgroup for “TOW service arrived” does not follow a normal distribution.

- **Step-4:** Conduct the estimation of clearance time for those groups with the classical multiple regression if they have sufficient samples and follow the pattern of normal distribution.
- **Step-5:** Apply the hazard-based modeling method [118, 119] for those groups with sufficient samples but without normally distributed clearance times.

Note that the hazard-based estimation method is commonly used to describe the data in the form of time from a well-defined time origin until the occurrence of some particular events of an end-point [118, 119]. It is well suited for use in developing predicting models for incident duration as reported in the literature [7, 17, 18, 21, 22, 27, 28, 31, 33]. Such an estimation model generally has two alternative forms to estimate the hazard function for a specific survival problem: the proportional hazard (PH) metric and the accelerated failure time (AFT) metric model [120]. The key assumption for an AFT model is that the survival time accelerates (or decelerates) by a constant factor when comparing different levels of covariates, whereas a PH model is grounded on the assumption of having a constant hazard ratio [21].

Based on their core assumptions, traffic safety researchers generally view the AFT model to be more appropriate to model the incident duration [7, 28]. Mathematically, one can present the AFT model with the following expression:

$$\log(T_i) = \beta_0 + \beta_1 x_{i,1} + \dots + \beta_p x_{i,p} + \sigma \varepsilon_i \quad (1)$$

where T_i is an incident duration, x_i is an explanatory variable, β_p is a corresponding parameter, σ is a scale parameter and ε_i is an error term. Depending on the distribution of the error terms, one can specify an AFT estimation method as a Weibull, log-logistic, or log-normal model to yield the best estimate.

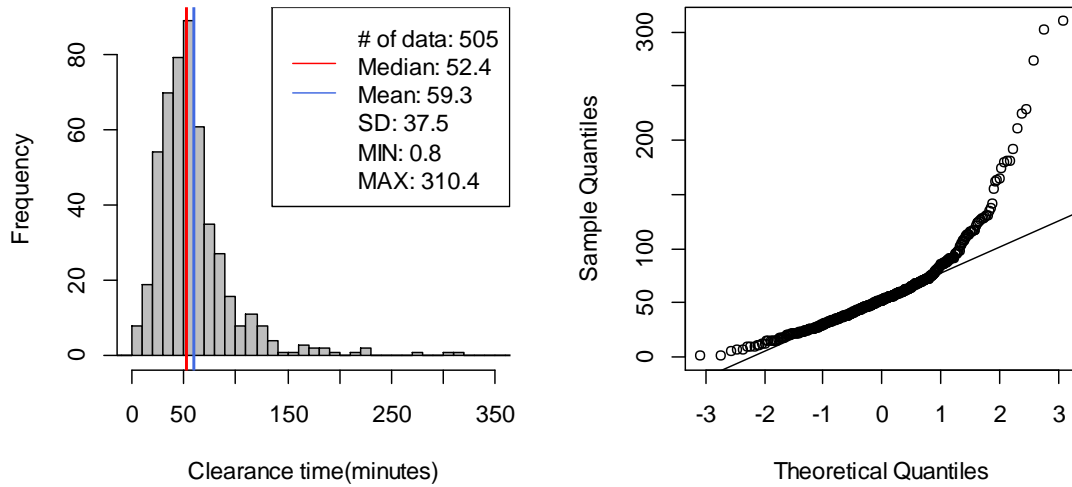


Figure 20. Distribution and Q-Q plot (e.g., subgroup for “TOW arrived” cases)

5.2.3. Stage-3: Ensemble of all estimates

Conceivably, one can have two estimates of clearance time for a detected incident with the above information, i.e., from the primary model and the sub-model to which the detected

incident belongs. The estimate from the former is to reflect the collective impacts of all related factors on the detected incident's clearance time, while the latter is to capture the dominating impacts imposed by one most critical factor (i.e., fatality) associated with that particular type of incidents. Those two estimates shall be combined to achieve more robust results, because 1) the generally small sample size with a small degree of freedom in each subgroup may degrade the quality of estimated parameters [121-123]; 2) some incidents can also be affected by more than one dominating factors; and 3) the development of a supplemental relation between these two estimates may best use the information for available data [124]. Hence, the final estimated clearance time (CT_i^e) for each detected incident shall be presented as the weighted combination of two estimates as follows:

$$CT_i^e = w_0 \cdot CT_0 + w_i \cdot CT_i \quad (2)$$

where CT_0 is the estimate from the primary model and CT_i is the one from its associated sub-model; w_0 and w_i are their respective weights.

Note that there are various methods available in the literature for estimating such weights. This study has adopted the robust regression method to produce the weights for each sub-model, because it can circumvent the limitations of non-normality and a tail risk without loss of information [125]. The estimate from each group is viewed as an independent variable for performing the robust regression, and the estimated coefficient for each independent variable is the weight (from w_0 to w_{n+1}) for each sub-model in Equation (2).

5.3. Estimation results

Table 6 presents the implementation results from modeling Stage-1 with respect to the dataset for 2009 incidents in the test set (years 2012-2015), including the dominating factor, median value, standard deviation, and available sample size for each classified group. The finally adopted estimation method at Stage-2 for each classified group and for the primary model with all data are also shown in Table 6.

Table 7 and Figure 21 show the estimation results for those groups and models classified in Table 6. Each modeling method and its distribution assumption have been chosen based on the statistical indicators such as R-squared value, Log-likelihood value, and Akaike's Information Criterion (AIC). The results with the rule-based modeling method for the groups of incidents having only small samples (e.g., fatality incidents, hazardous materials incidents) are shown in Figure 21.

Lastly, Table 8 shows the estimation results from Stage-3, i.e., using the robust regression with Huber weighting function [126] to compute the weights for each group since it has shown the best statistical results in this study. The coefficients of the robust regression model are used as the weights of Equation (2). Those estimates for incident clearance time with the ensemble model can be calculated from those weights and the estimates from each associated group. Note that data from the Year 2016 was used for validating of the rule-based models, developed for those incident groups having a small sample size.

Table 6. Summary of Classification Results with the Proposed Methodology

Key factor		Median (mins)	SD	# of cases	Applied modeling method
All cases		34	46	2009	AFT model (Weibull)
Sub group	Incident type: Fatality	236	97	21	Rule-based model*
	Truck over-turned	160	94	29	Multiple linear regression
	Truck lost-load	96	9	2	Directly using the median
	6+ lanes blocked	94	92	30	AFT model (Log normal)
	Hazard material related	74	80	9	Rule-based model*
	Vehicle jack-knifed	66	35	11	Rule-based model*
	Vehicle over-turned	59	35	96	AFT model (Weibull)
	Medical service arrived	53	42	35	AFT model (Log logistic)
	TOW service arrived	52	37	505	AFT model (Log logistic)
	County: Harford region	32	33	137	AFT model (Weibull)
	Incident type: Injury	30	23	345	AFT model (Weibull)
	Pavement condition: Wet	26	22	99	AFT model (Weibull)
	Truck involved	22	27	154	AFT model (Weibull)
	Night time	22	21	111	AFT model (Weibull)
	Center: AOC	20	20	334	AFT model (Weibull)
PM-peak hour	14	18	23	Directly using the median	
Remaining cases	9	18	68	AFT model (Weibull)	

* The detail process for the rule-based model is presented in [10].

Table 7. Model Estimation Results for Each Group

Variable	Estimate (p-value)												
	All cases	Truck over-turned	6+ lanes blocked	Vehicle over-turned	Medical service arrived	TOW service arrived	Harford region	Injury	Wet condition	Truck involved	Night time	AOC	Remaining case
	Weibull AFT model	Multiple linear reg.	Log-normal AFT model	Weibull AFT model	Log-logistic AFT model	Log-logistic AFT model	Weibull AFT model	Weibull AFT model	Weibull AFT model	Weibull AFT model	Weibull AFT model	Weibull AFT model	Weibull AFT model
(Intercept)	3.54(0.00)	12.94(0.82)	0.96(0.05)	3.34(0.00)	2.00(0.00)	2.99(0.00)	2.30(0.00)	2.61(0.00)	2.10(0.00)	2.14(0.00)	2.47(0.00)	2.23(0.00)	2.63(0.00)
Injury = 1	-0.97(0.00)		0.94(0.00)	0.21(0.08)			0.17(0.2)						
Property damage = 1	-1.04(0.00)												
Hazard material = 1	0.42(0.02)												
# of blocked lanes	0.08(0.00)	21.28(0.05)				0.05(0.07)	0.19(0.04)		0.23(0.05)		0.15(0.14)	0.28(0.00)	0.17(0.03)
# of blocked travel lanes				0.15(0.05)									
# of blocked auxiliary lanes	0.14(0.01)	51.37(0.16)							0.46(0.05)				
# of involved vehicles	0.04(0.01)				0.22(0.00)	0.05(0.01)	0.15(0.02)	0.13(0.00)			0.12(0.19)		
# of involved trucks	0.15(0.00)		0.50(0.00)			0.27(0.00)	0.55(0.00)						
Vehicle over-turned = 1	0.28(0.00)		0.43(0.01)										
Truck over-turned = 1	0.85(0.00)												
Truck lost-load = 1	0.81(0.10)												
# of response units	0.13(0.00)		0.16(0.00)			0.10(0.00)	0.17(0.01)	0.16(0.00)	0.18(0.01)	0.29(0.00)	0.19(0.02)	0.27(0.00)	
# of CHART													-0.37(0.11)
# of police		56.02(0.11)			0.85(0.01)								0.46(0.06)
# of fireboards			0.45(0.10)										-0.55(0.02)
# of medical services			0.37(0.16)										
# of TOW services	0.29(0.00)			0.18(0.02)	0.52(0.00)								
Weekend = 1		46.15(0.23)											
Daytime = 1			0.76(0.00)	0.45(0.00)				-0.10(0.18)					
PM-peak = 1			1.54(0.00)					-0.30(0.00)					
Night time = 1	0.13(0)		0.87(0.00)	0.42(0.00)	0.25(0.06)	0.19(0.00)	0.32(0.02)						
AOC = 1	0.35(0)					0.24(0.00)		0.45(0.00)	0.55(0.02)	0.72(0.00)			
TOC3 = 1											-1.03(0.00)		
TOC4 = 1											-0.92(0.00)		
Baltimore city = 1			0.45(0.01)										
Harford = 1	0.15(0)			0.25(0.04)	0.38(0.01)								
Cecil = 1					0.44(0.01)								
Log(scale)	-0.36(0)		-1.10(0.00)	-0.75(0.00)	-1.62(0.00)	-1.27(0.00)	-0.41(0.00)	-0.49(0.00)	-0.29(0.00)	-0.27(0.00)	-0.30(0.00)	-0.09(0.05)	-0.11(0.25)

* AFT: Accelerated Failure Time metric model

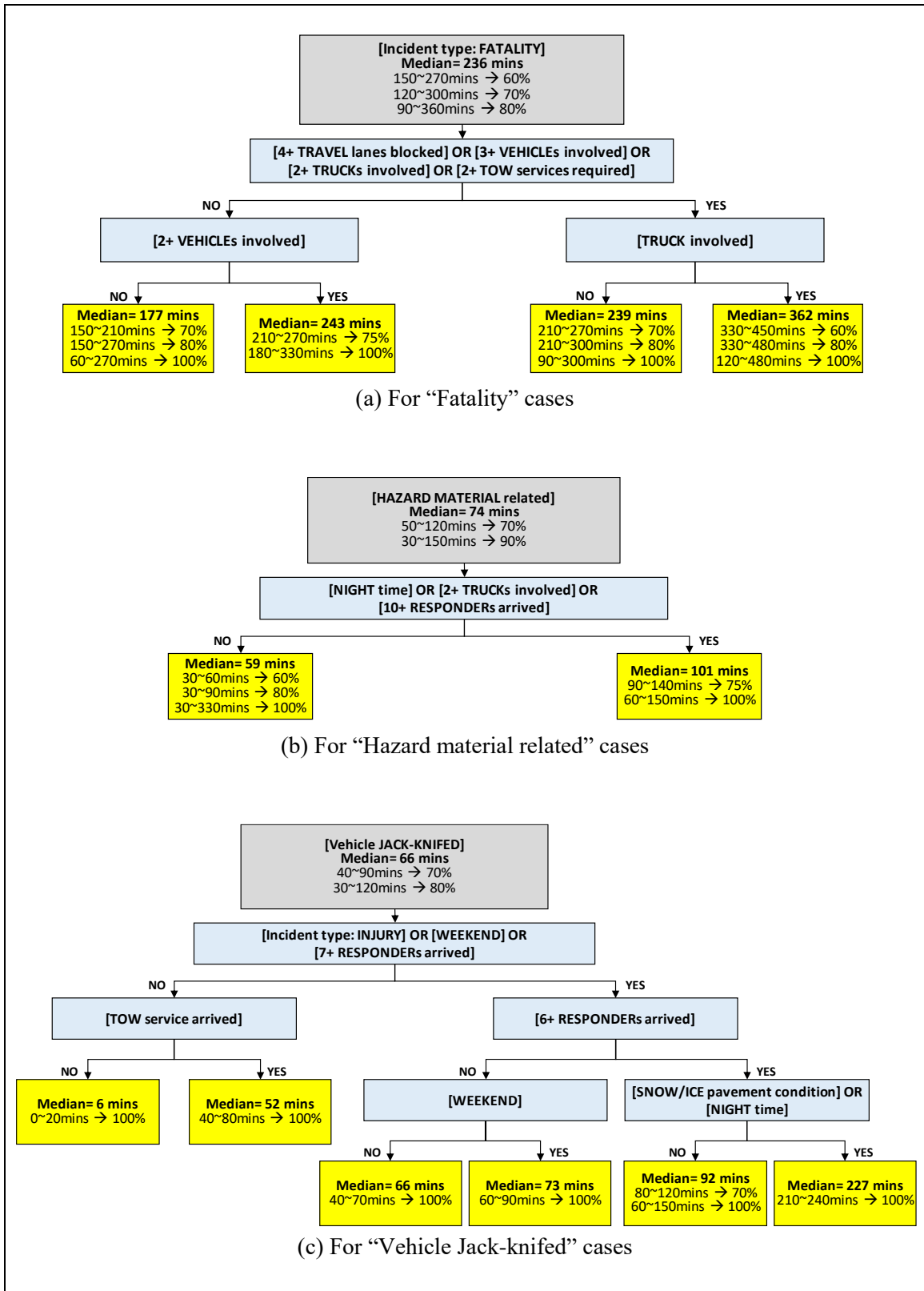


Figure 21. Rule-based models for incidents having a small sample size (i.e., fatality, hazard material, and vehicle jack-knifed cases)

Table 8. Estimated Weights for Each Group Produced from Robust Regression

Variable	Estimate	Std. Error	t-value
(Intercept)	-0.409	1.210	-0.338
All cases	0.100	0.013	7.811
Incident type: Fatality	0.850	0.025	34.049
Truck over-turned	0.818	0.029	28.182
Truck lost-load	0.895	0.143	6.239
6+ lanes blocked	0.929	0.025	37.490
Hazard material related	0.907	0.083	10.977
Vehicle jack-knifed	0.946	0.088	10.739
Vehicle over-turned	0.738	0.034	22.073
Medical service arrived	0.911	0.056	16.244
TOW service arrived	0.914	0.029	31.333
County: Harford region	0.734	0.041	17.943
Incident type: Injury	0.756	0.040	18.927
Pavement condition: Wet	0.763	0.070	10.890
Truck involved	0.788	0.057	13.789
Night time	0.762	0.074	10.366
Center: AOC	0.802	0.062	13.034
PM-peak hour	1.150	0.301	3.822
Remaining cases	0.819	0.151	5.435

Table 7 shows the estimation results for the primary and supplemental models for twelve types of incidents. From the number of significant variables in each supplemental model, one can evidence the need to properly decompose the entire incident dataset into several subgroups, and to select the most appropriate estimation method, based on the data distribution characteristics. For example, the duration for incidents involving over-turned trucks is mainly dominated by four key factors in the dataset, although some additional factors shown in the primary model may also play some roles but to less extent.

The results in Table 7 also confirm that one factor may play a different role on the clearance time of different incident types. For example, “the number of response units” is significant in most incident types, except for the incidents of “Truck over-turned”, “Vehicle

over-turned” and “Medical service arrived”. Similarly, “Weekend” is significant under the “Truck over-turned” condition, but not under the other incident types.

Lastly, Table 8 shows that the weighted estimates for each supplemental model when integrating with the primary model for all incident cases, reflecting the impacts of those factors, not included in each supplemental model, on the resulting incident duration of each incident type.

5.4. Model evaluation and discussion

To evaluate the performance of the developed continuous model, this section has selected the following three models for comparison:

- **Model-1:** AFT model developed with data from all cases (CT_0)
- **Model-2:** Only the set of individual models; each developed specially for each group (CT_i)
- **Model-3:** Ensemble model with the calibrated weights ($w_0 \cdot CT_0 + w_i \cdot CT_i$)

Model-1, serving as the baseline, is the AFT model using the information from all cases. Model-2 shows the set of individual models; each best fitting its own subgroup. Model-3 is the ensemble model with the calibrated weights for the AFT and each individual model. Table 9 shows the model accuracy using MAEs. Table 10 shows the model precision using the standard deviations of estimation errors.

Table 9. Accuracy (MAE) of point estimates by CT categories

	Model	Actual incident clearance time (minutes)				Overall
		< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	1	14.98	13.42	32.39	121.20	23.16
	2	12.86	12.21	27.57	74.38	18.54
	3	10.50	12.65	28.90	76.04	18.08
Test set (2017)	1	15.95	14.79	30.17	76.41	21.75
	2	14.86	13.11	29.22	67.47	20.01
	3	12.68	13.56	29.23	67.29	19.31

Table 10. Precision (SD of error) of point estimates by CT categories

	Model	Actual incident clearance time (minutes)				Overall
		< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	1	9.56	15.31	32.39	120.03	40.45
	2	9.81	10.98	23.20	57.45	23.59
	3	8.47	9.87	22.17	55.39	23.31
Test set (2017)	1	15.73	14.52	22.98	55.66	25.80
	2	12.77	10.99	20.03	44.18	21.63
	3	12.12	10.59	20.93	41.42	21.43

With respect to Model-2 developed from Stage-1 and Stage-2, it has reduced the MAEs from 23.16 to 18.54 and 21.75 to 20.01 in the training and test sets, respectively, compared with the baseline model. The same level of improvement also exists across all incident duration categories. Such improvements seem to assert our view that the required clearance time of an incident is often dominated by one or two most critical factors associated with incident nature. The overall precision of Model-2 also increases, as evident by reducing the standard deviation of estimation errors from 40.45 to 23.59 and 25.80 to 21.63 in the training and test sets, respectively. The improvement in precisions also exists across all incident duration categories, except for the category of less than 30 minutes. Overall, both the estimation accuracy and resulting precision, as expected, increase with the available sample size, as shown in Table 6. Note that the set of models with the proposed sequential modeling process shows the potential to improve the prediction

accuracy for the incidents with a longer duration but a small sample size. For example, with Model-2, the estimation result for those incidents exceeding 120 minutes show the reduction of MAE from 121.20 to 74.38.

Model-3 from Stage-3 also shows the performance improvement in the overall MAEs and standard deviations of the estimation errors (i.e., from 18.54 to 18.08 and 23.59 to 23.31, respectively), when compared with the results from the Model-2. The same pattern of improvement also shows in the test set (i.e., from 20.01 to 19.31 and 21.63 to 21.43 in accuracy and precision, respectively). Note that Model-3 seems to show consistent improvement in precision across all categories of incidents when compared with the results from Model-2 and Model-1. However, the performance of Model-3 with respect to MAEs seems to achieve the best results for incidents less than 30 minutes, significantly better than those by Model-2 and Model-1. However, Model-3, despite outperforming Model-1 in all categories of incidents, exhibits the comparable performance with Model-2, except for the category of incidents less than 30 minutes. Such results are attributable to the ensemble technique that can produce a better estimate by combining the strengths from different modeling results [124]; and the employed robust regression that can alleviate the estimation error due to some outliers on the tails [125].

In brief, the employed ensemble technique, coupled with the robust regression, can indeed improve the resulting accuracy and precision of the estimate by taking advantage of the unique strength offered by each sub-model, and is best for mitigating the bias caused by the tail risk in the data distribution.

Chapter 6. Supplemental models for outliers and integration

6.1. Identifying outliers among detected incidents

6.1.1. Modeling methodology

As mentioned previously, the distribution of incident duration data is highly right-skewed, exhibiting a long tail on its right side. Most of such datasets have the clearance duration of less than 60 minutes, and less than 10% of incidents exceeding 120 minutes. Hence, it is noticeable that such small percentage of recorded incident data may reflect the anomaly nature of some responded incidents or data-recording errors due to human-factor related issues. Most existing models, calibrated to concurrently account for the relationships between the highly skewed data and available factors, have often concluded such incidents as outliers, which are significantly different from the normal data [63].

Since an outlier can either be an anomaly or noise, where the former is referred to a special kind of outlier which may reflect some unique incident patterns that deserve the attention of potential users. The noise points are mostly related to data errors due to either the communications or human-factor issues. A reliable classification of those outlier data points from available incident records will enable the responsible agencies to not only improve the accuracy of clearance time estimation, but also to identify deficiencies in response to some special types of incidents, including the resource needs, missing coordination between different agencies, and better personnel training as well as the recording process.

Hence, this task aimed to develop an outlier analysis process, includes 1) the outlier score computation to identify anomalies and noises, and 2) a supplemental function to

enhance the prediction accuracy of the proposed primary models for estimating the incident clearance time.

6.1.2. Model development

The proposed methodology intends for use not only on detecting outliers, but also finding useful information related to their unique patterns so as to improve the performance of incident duration models. To do so, it is designed to the following stages of operations:

- **Stage-1:** Compute the outlier scores for each data point using ensemble modeling techniques;
- **Stage-2:** Set a threshold for identifying the outlier scores and notable outliers;
- **Stage-3:** Investigate the pattern of the outliers with a proposed rule-mining method, and classify the outliers as either anomalies or noises.

Figure 22 shows the modeling process of the proposed methodology and its key steps used to supplement the primary prediction models.

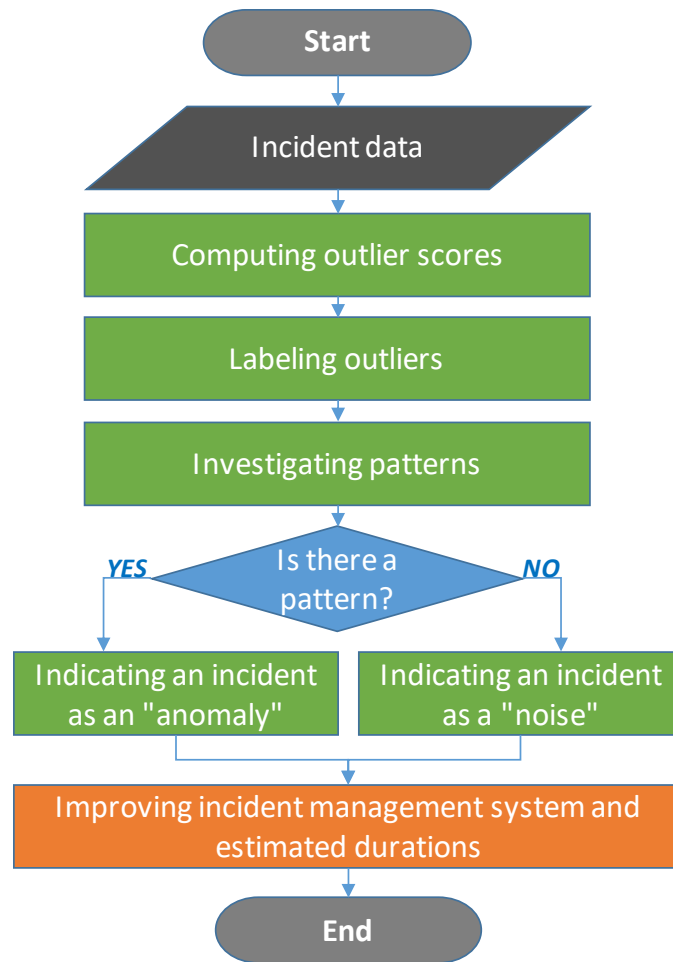


Figure 22. Flowchart of the modeling concept

6.1.2.1. Stage-1: Computation of the ensemble outlier scores

The first stage in the proposed methodology is to compute the outlier scores for each data point to rank them in the order of their outlier tendency [63]. Note that there are several types of outlier detection algorithms available in the literatures, including extreme-values, probabilistic/statistical models, linear models, and proximity-based models [63]. Each of those models only best fit to some datasets with unique characteristics or properties. In view of the large number of qualitative and quantitative variables associated with incident

durations and those complex correlations [10], it would be difficult to use only one of the existing outlier detection algorithms for various data patterns.

A possible solution is to use the ensemble technique which can combine two or more estimates from different algorithms to produce a robust result [63]. Hence, the following three algorithms with different strengths have been used at Stage-1 to yield the best performance: Principal Component Analysis (PCA) as a linear model, Partitioning Around Medoids (PAM) as a proximity-based model, and Isolation Forest for addressing the high dimensionality of the incident duration data. One can then compute the ensemble outlier scores, based on the results from each component.

6.1.2.1.1. Principal Component Analysis (PCA) as a linear model

PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables, known as principal components [127]. It is one of the dimension-reduction tools to reduce a large set of variables to a small set containing most of the information in the large set. In outlier analysis, PCA can be derived through multivariate regression analysis by determining the hyperplane that minimizes the least-squares error (i.e., distance) to the hyperplane [63]. The distances of the data points from this hyperplane, such as reconstruction errors, are used to quantify the outlier scores.

$$\text{Score}(\bar{X}) = \sum_{j=1}^d \frac{|(\bar{X} - \bar{\mu}) \cdot \bar{e}_j|^2}{\lambda_j} \quad (3)$$

The overall normalized outlier score is the sum of squares of a data point \bar{X} to the centroid $\bar{\mu}$ of the data, and \bar{e}_j is the j th eigenvector with a variance (eigenvalue) of λ_j along that direction in Equation (3). By using Equation (3), one can compute the outlier scores for each data point to show the distance of the data points from the normal data patterns along the lower-dimensional subspaces. Note that the PCA for outlier analysis usually provides promising results when there are significant correlations among the attributes. However, it is not suitable when the dataset exhibits an arbitrary shape or has a high dimensionality.

6.1.2.1.2. Partitioning Around Medoids (PAM) as a proximity-based model

The proximity-based methods are designed to model outliers as points that are isolated from the remaining data on the basis of similarity or distance functions [63]. The distance from a data point to its nearest centroid can be used to quantify the outlier scores. Such proximity-based methods can be effective when the datasets show arbitrary shapes [128].

Different from the K-means algorithm, PAM is a clustering algorithm which uses medoids instead of means in determining the optimal number of K clusters for the clustering process. Equation (4) shows the “silhouettes” function [128] used in the PAM to determine the “ K ” clusters.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (4)$$

For each datum i , $a(i)$ is the average distance between i and all other data within the same cluster, and $b(i)$ is the lowest average distance of i to all points in any other cluster.

With Equation (4), one can evaluate the quality of the clustered results and select the most appropriate number of clusters based on the value of $s(i)$, which can range from -1 to 1 where a higher value indicates a better quality of clustering. Note that the proximity-based methods are sensitive to noises in the dataset and do not perform well with the high-dimensional data.

6.1.2.1.3. Isolation Forest for high dimensionality

For high-dimensional data, the outlier scores from the irrelevant attributes can dilute the accuracy of some outlier detection algorithms, such as distance-based or density-based algorithms. For example, the data would become increasingly sparse, and all pairs of data points may be nearly equidistant from one another. Thus, the outlier scores would be less distinguishable from one another [129].

One of the promising alternatives is to use the Isolation Forest algorithm [130] which can recursively partition the dataset with axis-parallel cuts at randomly chosen partition points from randomly selected attributes. By doing so, it can isolate the instances into nodes with fewer and fewer instances until all points are isolated into singleton nodes, each containing only one instance [63]. The tree branches containing outliers are noticeably less deep, because these data points are located in sparse regions. Thus, the distance of the leaf to the root is used as the outlier score [63].

Isolation Forest is a subspace outlier detection algorithm that can be used to identify hidden outliers, which are masked by high dimensionality, in the low-dimensional subspaces. Therefore, isolated outliers can be effectively identified in the multidimensional space. However, the Isolation Forest algorithm is sensitive to those noises in the dataset,

and the scores from Isolation Forest do not indicate the intensity of outlierness. This means that the outlier scores from Isolation Forest represent the hierarchies of outlierness, whereas the outlier scores from PCA and PAM denote the distance of a data point from the normal cases. Therefore, the Isolation Forest algorithm can underestimate the intensity of outlierness.

6.1.2.1.4. Ensemble outlier scores

All three outlier detection methods described in the previous sections have their strengths and weaknesses. More specifically, PAM and Isolation Forest are sensitive to the noises in the dataset, but not the PCA which is effective for detecting outliers that have significant correlations among their attributes, but it is not appropriate for detecting the outliers in arbitrary shapes. Furthermore, PAM is appropriate for detecting the outliers in arbitrary shapes, but it may miss some of the outliers in the high-dimensional spaces. Isolation Forest is a dimensionality reduction process that can effectively identify the hidden outliers isolated in the high-dimensional spaces, but it does not reflect well the intensity of outlierness. Therefore, to take advantage of those outlier detection algorithms, but circumvent their deficiencies, this study proposes to normalize the outlier score from each of those three algorithms and combine them with the following equation:

$$\text{Ensemble Score}(\bar{X}) = \sum_{k=1}^n \frac{\bar{X}_k - \bar{X}_k^{\min}}{n \cdot (\bar{X}_k^{\max} - \bar{X}_k^{\min})} \quad (5)$$

In Equation (5), n is the number of outlier detection components k ; \bar{X}_k^{max} and \bar{X}_k^{min} denote the maximum and minimum outlier scores, respectively, for each component. By using Equation (5), one can normalize each outlier score and compute the ensemble outlier scores that lie between the minimum of “0” and the maximum value of “1”, respectively.

6.1.2.2. Stage-2: Outlier labeling

The next step is to label the outliers using the ensemble outlier scores, which allows the user to determine those data points that shall be classified as an outlier, based on a specified threshold. Although the outlier scores are useful for evaluating the outlierness of each data point, one needs to further distinguish whether a data point is a notable outlier or data error.

There are two types of information for labeling the outlier: 1) the dispersion of outlier scores, and 2) the relation between the outlier scores and the incident durations. This is due to the fact that a notable outlier generally has a large outlier score, and its duration is most likely distributed in the range of its right-skewed tail of the distribution.

6.1.2.3. Stage-3: Pattern examination and classification

The task at the final stage is to investigate the pattern of the outliers and divide the outliers into anomalies and noises with the proposed following steps:

- **Step-1:** From the training set of outliers (G_T), search for a classification rule to discriminate the selected outlier ($O_{i=1}$, an outlier with the longest incident clearance time) using the association rule mining method [114, 115].

- **Step-2:** Select a critical rule with approximately more than 75% confidence level and the highest support level, and validate whether the selected rule also has the same confidence level in the combined group ($G_T + G_V$).
- **Step-3:** If there is no critical rule or the selected rule does not fulfill the condition of Step-2, include one more outlier (O_{i+1}) and go to Step-1, and then search for a classification rule again.
- **Step-4:** Compute the mean and range of incident clearance times of the selected outliers associated with the classification rule, and filter them out from the training and test sets (G_T and G_V).
- **Step-5:** Stop the process if there is no rule or no remaining data in the test set.
- **Step-6:** Otherwise, go to Step-1 and repeat the same process for the remaining outliers.

G_T and G_V denote the training and validation sets of an outlier data group, respectively; O is the selected outlier from each data set; i is the number of selected outlier with the longest duration; therefore, $O_{i=1}$ shows one outlier with the longest duration and $O_{i=2}$ denotes two outliers with the first and second longest duration, and so on.

A key feature of the proposed mining method with the hybrid association rules is that one can preferentially identify the pattern of outliers with long duration in the highly imbalanced classification problem [10]. Using this proposed algorithm, one can also divide the outliers into anomalies and noises, depending on whether they can be classified by the rules or not. The logic behind this idea is that any outlier without any explainable pattern can be classified as a noise.

6.1.3. Estimation results

6.1.3.1. Stage-1: Computation of the ensemble outlier scores

The first computation for the outlier scores includes only continuous variables, such as the number of blocked lanes, the number of involved vehicles, and the number of response units, because the normality assumption underlying PCA will be violated if a variable is discrete [131]. Table 11 shows the results of the PCA. Since its cumulative proportion of the fifth principal component is around 0.95, one shall take the first five principal components to compute the outlier scores with Equation (3). Figure 24 (a) shows the distribution of these outlier scores from the test set.

Table 11. Estimation Results of the PCA

		PC1	PC2	PC3	PC4	PC5	PC6	PC7
Variable	# of blocked travel lanes	-0.145	-0.106	0.048	-0.861	-0.326	0.340	-0.042
	# of blocked shoulders	-0.078	-0.022	0.081	-0.445	0.772	-0.438	-0.015
	# of involved vehicles	-0.041	-0.992	-0.027	0.114	0.033	-0.004	-0.017
	# of involved trucks	-0.050	-0.025	0.001	-0.112	-0.542	-0.831	-0.003
	# of arrived total responders	-0.911	0.044	0.236	0.159	0.000	0.031	0.293
	# of arrived CHART units	-0.332	0.044	-0.883	-0.004	0.046	-0.011	-0.325
	# of arrived RES_TOW	-0.168	0.023	0.393	0.100	-0.013	0.009	-0.898
Standard deviation		1.974	1.062	0.732	0.687	0.508	0.442	0.409
Proportion of Variance		0.586	0.169	0.081	0.071	0.039	0.029	0.025
Cumulative Proportion		0.586	0.755	0.836	0.907	0.946	0.975	1.000

Figure 23 shows the silhouette values according to the number of clusters (k), where one shall take $k = 2$ since it corresponds with the highest silhouette value. Figure 24 (b) shows the distribution of these outlier scores in the test set.

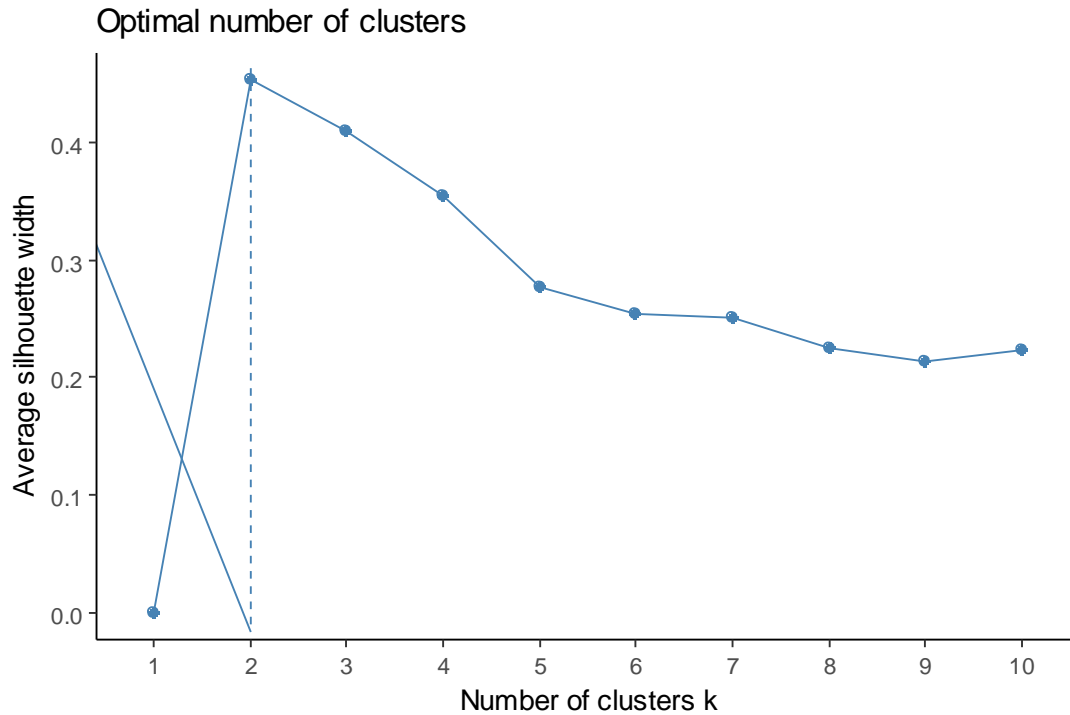


Figure 23. Silhouette values to determine the optimal number of clusters for PAM

For applying Isolation Forest, those variables involved for computation are: lane blockage, involved vehicle, responder, incident type, environmental conditions, operation center, time of a day, and location information. A total of 300 sub-samples and 200 trees were used for executing the Isolation Forest algorithm [130]. Based on the results of Isolation Forest, one can then compute the path lengths for each data point from the root to the leaf, and these path lengths can be used as the outlier scores. Figure 24 (c) shows the distribution of these inverted outlier scores in the test set.

Lastly, Figure 24 (d) shows the distribution of these ensemble outlier scores based on three different algorithms and Equation (5) in the test set. Table 12 shows its quantile values of the computing results.

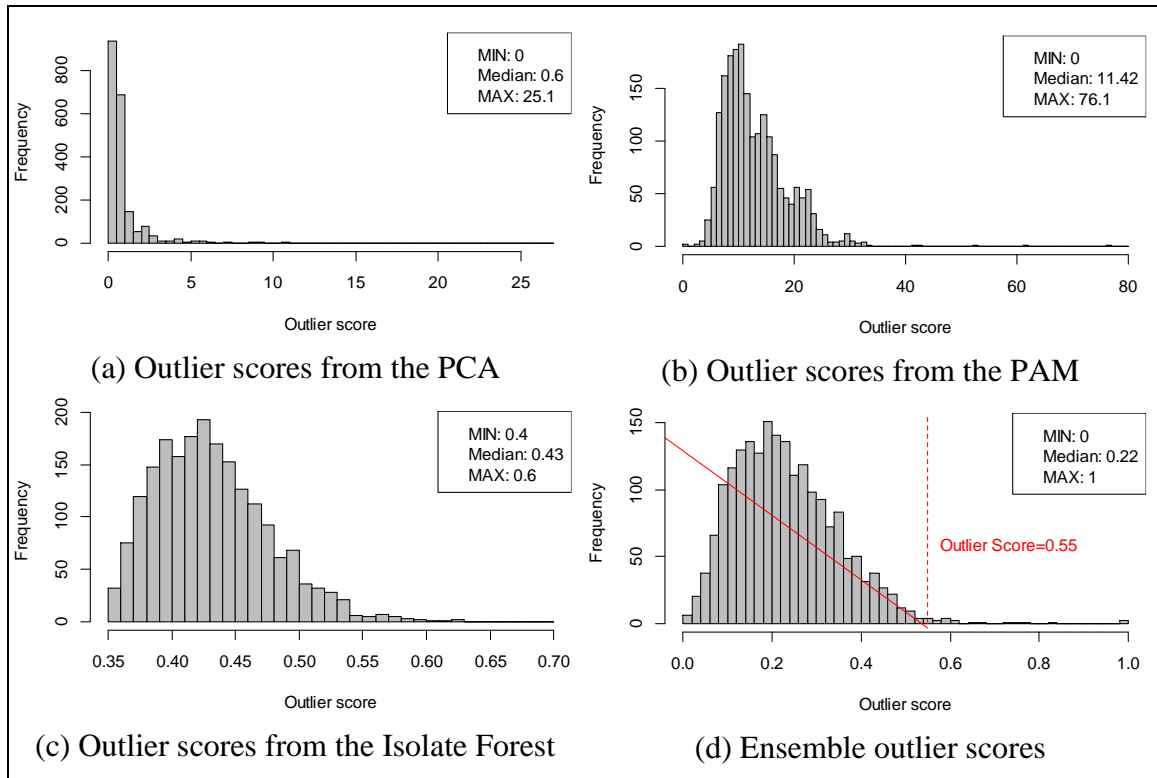


Figure 24. Distributions of the outlier scores for each outlier detection algorithm

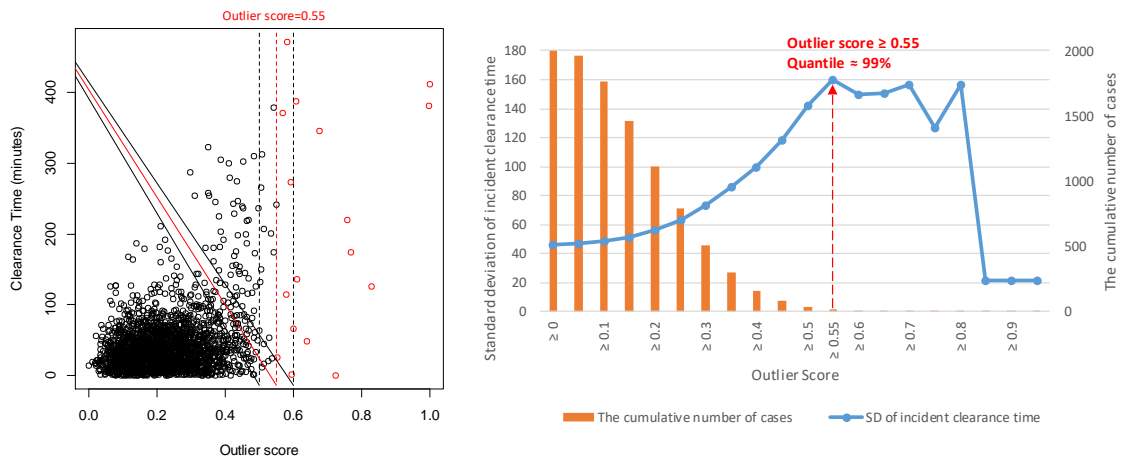
Table 12. Quantile values for the ensemble outlier scores

Quantile	10%	30%	50%	70%	90%	95%	97%	98%	99%
Outlier score	0.094	0.158	0.216	0.281	0.383	0.436	0.463	0.484	0.537

6.1.3.2. Stage-2: Outlier labeling

Figure 24 (d) shows the distribution of the outlier scores for each data point, and Table 12 shows its quantile values. Figure 25 (a) shows the relations between the outlier scores and incident clearance times. Figure 25 (b) shows the distribution of the incident clearance times and the cumulative number of cases larger than a specified outlier score. As shown in Figure 25 (a), the incident clearance times show a different pattern when the outlier scores are higher than 0.55. Also, it is noticeable that the incident data are very sparse when

the outlier scores are higher than 0.55, as shown in Figure 25 (b). Lastly, since 99% of the quantile value is a reasonable threshold to distinguish the outliers from the normal cases (see Figure 24 (d)), this study has set the threshold to be 0.55.



(a) Relation between outlier scores and incident clearance times

(b) Standard deviation for the incident clearance time and the cumulative number of the cases according to the outlier scores

Figure 25. Outlier score information to determine the threshold of outliers

6.1.3.3. Stage-3: Pattern examination and classification

The task at the final stage is to employ the rule mining method to determine the unique patterns of the anomalies and identify the noises. Note that Figure 26 shows the classification rules and the identified all outlier patterns, which reveal valuable information for improving the estimation accuracy of incident duration prediction.

If Outlier Score ≥ 0.55			
Description		CT range (minutes)	Mean CT (minutes)
IF [Fatality] & [12+ response units arrived]	THEN	390~480	442
IF [5+ lanes blocked]	THEN	270~390	352
IF [Fatality]	THEN	180~240	220
IF [2+ lanes blocked] & [7+ response units arrived]	THEN	120~240	150
IF [2+ TOW services arrived]	THEN	90~150	126
IF [Truck over-turned]	THEN	60~120	66
IF [3+ response units arrived] & [1+ TOW services arrived]	THEN	30~90	48
ELSE THEN		"Noise"	

Figure 26. Classification rules for the outliers

6.1.4. Model evaluation and discussion

One of the objectives of this study is to improve the accuracy of incident duration estimation using outlier information. To assess the contributions of those supplemental rules from the outlier analysis, both models proposed in the previous chapters are used as the base model for predicting incident duration. Tables 13-15 show the evaluation results in the training and test sets. Note that the dataset of Year 2016 was used for validation to circumvent the overfitting issue. The AFT model (i.e., hazard-based model), which is one of the most popular in the literature, was selected as the baseline to assess the improvement.

Table 13. Accuracy (percentage) of interval estimates by CT categories

Dataset	Model	Actual CT (minutes)				overall
		< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	Classification model	96.1%	81.6%	35.7%	47.4%	77.2%
	+ Rules for outliers	96.1%	81.6%	35.7%	47.4%	77.2%
Test set (2017)	Classification model	97.0%	83.2%	33.8%	60.0%	77.9%
	+ Rules for outliers	97.0%	83.2%	35.0%	60.0%	78.2%

Table 14. Accuracy (MAE) of point estimates by CT categories

Dataset	Model	Actual CT (minutes)				overall
		< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	<i>Base line: AFT model</i>	14.98	13.42	32.39	121.20	23.16
	Continuous model	10.50	12.65	28.90	76.04	18.08
	+ Rules for outliers	10.50	12.64	28.45	68.23	17.59
Test set (2017)	<i>Base line: AFT model</i>	15.95	14.79	30.17	76.41	21.75
	Continuous model	12.68	13.56	29.23	67.29	19.31
	+ Rules for outliers	12.68	13.56	28.98	60.84	18.89

Table 15. Precision (SD of error) of point estimates by CT categories

Dataset	Model	Actual CT (minutes)				overall
		< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	<i>Base line: AFT model</i>	9.56	15.31	32.39	120.03	40.45
	Continuous model	8.47	9.87	22.17	55.39	23.31
	+ Rules for outliers	8.47	9.88	21.69	49.78	21.52
Test set (2017)	<i>Base line: AFT model</i>	15.73	14.52	22.98	55.66	25.80
	Continuous model	12.12	10.59	20.93	41.42	21.43
	+ Rules for outliers	12.12	10.59	20.54	38.87	20.24

The evaluation results shown in those tables clearly indicate that the performances of both classification and continuous models with respect to the incidents of more than 60 minutes increase with the use of those supplemental rules. For example, the accuracy shown in Table 13 slightly increases from 33.8% to 35% for incidents between 60 to 120 minutes in the test set.

The improvement for the continuous model seems more pronounced where both the MAE and standard deviation of errors decrease from 18.08 to 17.59 and 23.31 to 21.52, respectively. The same pattern of improvement exists across all incident duration categories, but most significantly in the category of “≥120” minutes, due to the nature of those supplemental rules. The overall accuracy for the test set has also been improved, as evidenced from the decreasing MAE and standard deviation of errors (i.e., from 19.31 to 18.89 and 21.43 to 20.24 respectively).

Table 16. Improvement of point and interval estimates after the rules of outliers

No.	Date	Actual CT (mins)	Point estimate (estimated CT – actual CT)		Interval estimate			
			Before	After	Before (with different confidence level)			After (100% confidence level)
					Around 60%	Around 70%	Around 80%	
1	08/16/2012	472	418 (-53)	442 (-30)	330~450	330~480	120~480	390~480
2	06/25/2015	411	407 (-4)	442 (31)	330~450	330~480	120~480	390~480
3	12/13/2016	406	468 (63)	442 (36)	330~450	330~480	120~480	390~480
4	10/15/2014	387	264 (-123)	352 (-35)	≥ 170	≥ 140	≥ 120	270~390
5	08/20/2013	382	262 (-120)	352 (-30)			≥ 120	270~390
6	06/20/2013	372	324 (-48)	352 (-20)	≥ 170	≥ 140	≥ 120	270~390
7	11/17/2013	346	388 (41)	352 (6)	330~450	330~480	120~480	270~390
8	10/29/2016	313	145 (-168)	352 (39)	60~120	45~120	25~120	270~390
9	12/06/2013	274	603 (329)	352 (78)	≥ 170	≥ 140	≥ 120	270~390
10	03/31/2014	220	347 (127)	220 (0)	330~450	330~480	120~480	180~240
11	07/20/2017	193	350 (156)	220 (27)	330~450	330~480	120~480	180~240
12	12/17/2016	227	224 (-2)	150 (-77)		≥ 120	≥ 100	120~240
13	03/10/2014	174	145 (-29)	150 (-24)		≥ 170	≥ 120	120~240
14	04/26/2017	136	182 (45)	150 (14)	≥ 240	≥ 120	≥ 60	120~240
15	12/22/2014	126	240 (114)	150 (24)	≥ 240	≥ 120	≥ 60	120~240
16	10/11/2015	136	193 (56)	126 (-10)	≥ 170	≥ 140	≥ 120	90~150
17	03/01/2012	115	162 (47)	126 (11)		≥ 120	≥ 100	90~150
18	02/12/2017	117	143 (26)	66 (-51)	20~60	10~60	10~70	60~120
19	02/12/2017	92	163 (71)	66 (-26)	60~120	45~120	25~120	60~120
20	02/15/2014	66	184 (118)	66 (0)	20~60	10~60	10~70	60~120
21	04/09/2016	77	87 (10)	48 (-29)	60~120	25~120	15~120	30~90
22	11/30/2016	58	56 (-1)	48 (-10)	20~60	10~60	10~70	30~90
23	09/06/2016	49	68 (19)	48 (-1)	25~60	15~60	10~60	30~90
24	06/03/2013	48	51 (3)	48 (0)	25~60	15~60	10~60	30~90
25	12/02/2016	36	88 (52)	48 (12)	60~120	55~120	20~120	30~90

* The number in the parenthesis indicates an estimation error (estimated CT – actual CT); the yellow color indicates the improved results in the point estimates and correctly classified cases in the interval estimates.

The results of further investigation with respect to those identified 25 anomalies (out of 29 outliers) are shown in Table 16. Among 29 outliers, 25 cases are detected as anomalies, and 4 cases are suspected of noises. Those rules developed from the outlier analysis improve the 20 estimation results out of 25 anomalies, and the MAE of all 25 anomalies decreases from 73.08 to 24.86. Two anomalies misclassified by the interval-based model are correctly estimated by the rules of outliers. Such improvements seem to

confirm the potential of using outlier information to improve incident duration estimates with either classification or continuous models as long as the dataset exhibits a long tail and skewed pattern.

6.2. An integrated estimation process

6.2.1. Modeling concept and methodology

The last component of the proposed incident duration model is the adjustment process when the interval and point estimates for one incident scenario yield inconsistent results. There can be two possible patterns when a point estimate is out of the boundaries of an interval estimate, as shown in Figure 27.

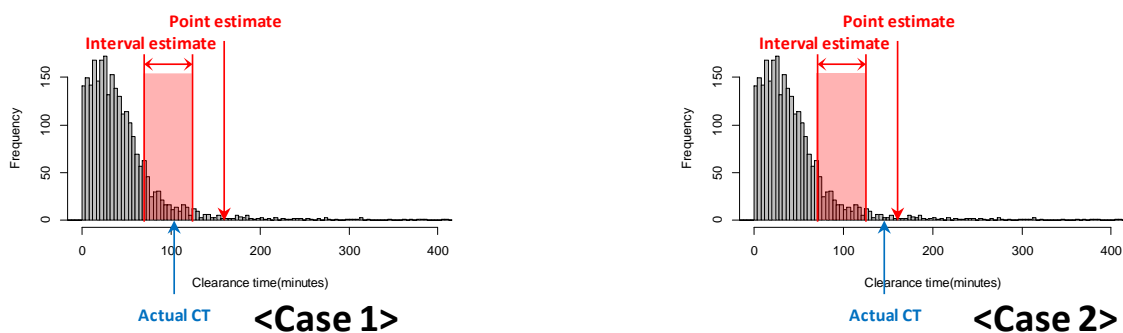


Figure 27. Examples when a point estimate and an interval estimate are mismatched

The first possible pattern is when the actual incident clearance time is in the boundaries of the estimated interval, and the second is when the actual incident clearance time is out of the boundaries of the estimated interval. Hence, one shall find some way to adjust such estimates to ensure their consistency.

6.2.2. Adjustment for inconsistent estimates

Depending on the position on the actual incident clearance time, either the interval or point estimates can be adjusted with the following process:

- **Step-1:** Move the point estimate to the adjacent boundary of the interval estimate, when the actual incident clearance time is within the boundaries of the interval estimate.
- **Step-2:** Extend the adjacent boundaries of the interval estimate to the point estimate, when the actual incident clearance time is out of the boundaries of the interval estimate.

Based on this process, one can make a decision tree model to adjust those inconsistent estimation results to provide more accurate results.

6.2.3. Estimation results

To identify which estimate result should be adjusted, the decision tree model (CART algorithm) is applied in this study. This study used the data of the years 2012-2015 as a training set and used the data of the year 2016 for model validation. Figure 28 shows the results using the decision tree to adjust inconsistent estimates.

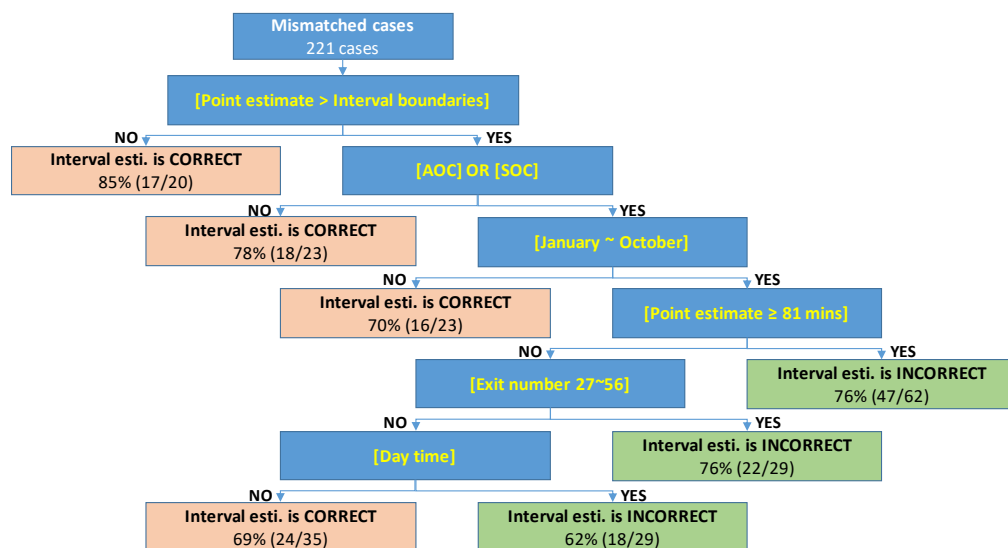


Figure 28. Decision tree model to adjust mismatched results (test set)

As seen in Figure 28, the developed decision tree with the integration process has achieved the accuracy level of exceeding 70%, after performing the adjustment for about 8% of the incident cases that show the inconsistent results.

6.2.4. Model evaluation and discussion

Table 17 shows the accuracy of interval estimates; Table 18 shows the accuracy of point estimates using MAE; and Table 19 shows precision of the point estimates based on the standard deviation of estimation errors. Note that the AFT model (i.e., hazard-based model), which is one of the most popular in the literature, was selected as the baseline to assess the improvement of the point estimates.

Table 17. Accuracy (percentage) of interval estimates by CT categories

Dataset	Model	Actual CT (minutes)				overall
		< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	Classification model	96.1%	81.6%	35.7%	47.4%	77.2%
	+ Rules for outliers	96.1%	81.6%	35.7%	47.4%	77.2%
	+ Rules for combining	96.1%	81.7%	38.0%	48.5%	77.7%
Test set (2017)	Classification model	97.0%	83.2%	33.8%	60.0%	77.9%
	+ Rules for outliers	97.0%	83.2%	35.0%	60.0%	78.2%
	+ Rules for combining	97.0%	83.2%	35.0%	64.0%	78.4%

Table 18. Accuracy (MAE) of point estimates by CT categories

Dataset	Model	Actual CT (minutes)				overall
		< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	<i>Base line: AFT model</i>	14.98	13.42	32.39	121.20	23.16
	Continuous model	10.50	12.65	28.90	76.04	18.08
	+ Rules for outliers	10.50	12.64	28.45	68.23	17.59
	+ Rules for combining	10.44	12.27	28.48	67.19	17.38
Test set (2017)	<i>Base line: AFT model</i>	15.95	14.79	30.17	76.41	21.75
	Continuous model	12.68	13.56	29.23	67.29	19.31
	+ Rules for outliers	12.68	13.56	28.98	60.84	18.89
	+ Rules for combining	12.43	13.28	29.05	59.38	18.62

Table 19. Precision (SD of error) of point estimates by CT categories

Dataset	Model	Actual CT (minutes)				overall
		< 30	30~60	60~120	≥ 120	
Training set (2012~2015)	<i>Base line: AFT model</i>	9.56	15.31	32.39	120.03	40.45
	Continuous model	8.47	9.87	22.17	55.39	23.31
	+ Rules for outliers	8.47	9.88	21.69	49.78	21.52
	+ Rules for combining	8.40	9.40	21.54	49.49	21.28
Test set (2017)	<i>Base line: AFT model</i>	15.73	14.52	22.98	55.66	25.80
	Continuous model	12.12	10.59	20.93	41.42	21.43
	+ Rules for outliers	12.12	10.59	20.54	38.87	20.24
	+ Rules for combining	11.37	10.39	20.87	41.66	20.30

In Table 17 for interval estimates, the accuracies of the interval-based results have slightly increased from 77.2% to 77.7% and 78.2% to 78.4% in the training and test sets respectively, and all cases can have consistent results for all categories of incident clearance time without any overfitting and performance reduction in the training set and the test set.

In Table 18 and Table 19 for point estimates, the overall accuracy and precision in the training set are slightly improved after the adjustment process, as evident in its decreased MAE and standard deviation of errors from 17.59 to 17.38 and 21.52 to 21.28, respectively. The overall accuracy is also improved, as reflected in the reduced MAE (from 18.89 to 18.62) in the test set. These results confirm that the combination process and its classification rules for the mismatched cases can indeed improve the estimation results without overfitting even with the small sample size of anomaly data points.

Table 20. Acceptable accuracy of point estimates by CT categories

Dataset	Actual incident clearance time with different error term (minutes)												Overall		
	< 30			30~60			60~120			≥ 120					
	<±15	<±30	<±60	<±15	<±30	<±60	<±15	<±30	<±60	<±30	<±60	<±120	<±15	<±30	<±60
Training set (2012~2015)	77%	97%	100%	67%	97%	100%	31%	59%	93%	31%	48%	89%	62%	87%	96%
Test set (2017)	68%	94%	99%	61%	93%	100%	32%	59%	91%	36%	56%	96%	55%	84%	96%

Table 21. Distribution of point estimation errors by CT categories (Training set)

Errors		Actual CT (minutes)								Overall	
		< 30		30 ~ 60		60 ~ 120		≥ 120			
		Before*	After**	Before	After	Before	After	Before	After	Before	After
Overestimated	> 120	0	0	2	0	6	1	13	1	21	2
	60~120	4	1	13	2	12	3	10	1	39	7
	30~60	44	18	41	9	15	6	6	4	106	37
	15~30	295	154	65	55	25	10	1	4	386	223
	10~15	161	127	54	46	3	8	1	1	219	182
Within -10~10		244	434	339	313	65	73	5	7	653	827
Underestimated	-15~-10	0	14	84	95	27	25	2	3	113	137
	-30~-15	0	0	72	142	74	87	4	15	150	244
	-60~-30	0	0	3	11	94	109	7	13	104	133
	-120~-60	0	0	0	0	21	20	35	38	56	58
	< -120	0	0	0	0	0	0	13	10	13	10
Total # of cases		748		673		342		97		1860	
Acceptable Accuracy		54.1%	76.9%	91.2%	96.7%	88.6%	93.0%	73.2%	88.7%	81.8%	86.7%

* Before: AFT (i.e., Hazard-base model) model for comparison

** After: The developed incident duration models (point-based model + supplemental models)

*** Yellow area: Acceptable ranges of the error term

Table 22. Distribution of point estimation errors by CT categories (Test set)

Errors		Actual CT (minutes)								Overall	
		< 30		30 ~ 60		60 ~ 120		≥ 120			
		Before*	After**	Before	After	Before	After	Before	After	Before	After
Overestimated	> 120	1	0	0	0	0	0	3	0	4	0
	60~120	0	1	3	0	3	0	1	0	7	1
	30~60	12	9	11	6	4	0	3	1	30	16
	15~30	63	40	25	20	3	3	2	2	93	65
	10~15	35	23	11	13	4	1	1	2	51	39
Within -10~10		55	87	72	74	19	14	1	0	147	175
Underestimated	-15~-10	0	3	13	8	1	11	0	0	14	22
	-30~-15	0	3	19	29	19	18	0	5	38	55
	-60~-30	0	0	1	5	23	26	3	4	27	35
	-120~-60	0	0	0	0	4	7	11	10	15	17
	< -120	0	0	0	0	0	0	0	1	0	1
Total # of cases		166		155		80		25		426	
Acceptable Accuracy		54.2%	68.1%	90.3%	92.9%	91.3%	91.3%	88.0%	96.0%	80.5%	83.6%

* Before: AFT (i.e., Hazard-base model) model for comparison

** After: The developed incident duration models (point-based model + supplemental models)

*** Yellow area: Acceptable ranges of the error term

With respect to the acceptable accuracy for incident management, there is no generally acceptable standard. Because a widely accepted model is not available for use in practice yet and those tolerance ranges may vary with each agency's available resources, man-

power, and the nature of incidents. Nevertheless, most states at the experimental stage view the standard deviation of the durations for incidents under nearly identical scenarios as the acceptable range of variation in prediction. Therefore, this study has adopted the interval of ± 30 minutes as the acceptable range, because the average of deviations of incidents in Maryland SHA generally lies between 30–50 minutes [132]. And, this study also has used the different acceptable ranges for each category close to their lower boundaries to show the different effects of the developed model under different conditions.

In this context, as seen in Table 20, with respect to the acceptable accuracy for field operations of incident management, the proposed model can yield about 77% and 68% of accuracy for the estimates, respectively in the training and test sets within the interval of less than 15 minutes' deviation for those incidents of lasting less than 30 minutes. Also, about 97% and 93% of its estimates for incidents between “30~60” minutes are within the estimation error of 30 minutes, and about 93% and 91% estimates for those within the “60~120” minutes categories within the 60 minutes' deviation.

Tables 21 and 22 also show the distribution of the point estimation errors by different incident clearance time categories in the training and test sets, respectively. The AFT model ('Before'), a state-of-the-art method in the literature, was selected as the baseline to assess the improvement in the estimation results of the incident duration module ('After') developed in this study. The developed model can improve the prediction accuracy of many overestimated cases in each CT category, and overall 51 and 10 highly overestimated cases having longer than 60 minutes errors are improved in the training and test sets, respectively. Although those performance improvements are not consistent across the underestimated

cases in different categories due to the uneven distribution of samples, the resulting reduction in the distribution of the estimation errors for most categories is promising.

To sum up, the proposed model with the ensemble technique and supplemental models can achieve between 83.6% and 86.5% of the overall accuracy if the 30-minutes deviation is viewed as acceptable from the field operation's perspective. The proposed model's capability to yield around 85% of correct estimates for incident clearance time offers the potential for its use in field incident impact assessment and traffic management.

Chapter 7. Capacity analysis for a lane-closure highway segment: single-lane-open scenarios

7.1. Introduction

Highway incidents such as disabled vehicles and accidents are the major causes of traffic congestion in urban areas due to their resulting lane closure on the roadway. The capacity reduction at the lane-closure location will inevitably cause traffic conflicts and queue formation as well as propagation to its upstream segment [95]. To mitigate such impacts from the perspective of both traffic safety and efficiency, it is imperative that responsible highway agencies can have a reliable tool for estimating the expected reduction in roadway capacity, and thereby deploying proper control strategies to contend with the resulting traffic queues and congestion.

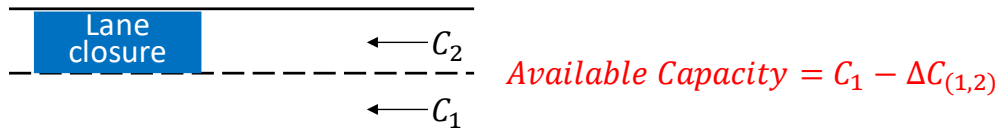


Figure 29. Example of the capacity drop in one-out-of-two-lanes-blockage scenario

Note that the lane closure by an incident often increases the number of traffic conflicts due to lane-changing maneuvers, and the outflows from the lane drop spot is generally lower than the expected capacity of the open lane [95, 96]. Figure 29 shows an example of the capacity drop phenomenon in the one-out-of-two-lanes-blockage scenario, where the actual available capacity for the outflows is, $C_1 - \Delta C_{(1,2)}$, less than the capacity of the open lane C_1 . The additional capacity reduction (i.e., $\Delta C_{(1,2)}$) at the lane-closure spot due to lane-changing flows is termed as ‘capacity drop’, which varies with various factors, including lane closure configuration, merging ratio, merging time and speed, acceleration rate of

vehicles in the traffic flows, etc. Hence, developing a reliable yet convenient model for estimating the capacity drop on a highway segment under different lane-closure scenarios is the focus of the chapter.

To estimate the capacity drop due to the lane closure, one can typically take the difference between the free-flow rate and actual queue discharge rate from the closed segment [96]. As described in the literature review, a large body of studies has been conducted for analyzing such capacity drop under various conditions, including work zones, ramps, lane closures, etc. Most of such studies apply one of the following analysis methods: statistical approaches, simulation-based approximation, and analytical models.

The statistical approaches for capacity analysis often suffer from having enough quality field data, such as various essential traffic data and lane-configuration conditions from the incident scenes. Besides, some key contributing factors may not be directly collectable from those incident sites. Moreover, those data-driven statistical approaches lack the proper functional form to reflect the physical dynamics collectively manifested by some key contributing factors, or to track high-order interactions between them.

With respect to the simulation-based microscopic approaches, their potentials for use in practice hinge on the availability of a reliable microscopic car-following and lane-changing models that can mimic traffic conditions during the lane-closure scenarios. Development of such models, however, remains an ongoing research task in the traffic community. Hence, to estimate the traffic impacts and queue length due to lane-closure incidents and also to maintain the trackability of such relations in real-time operations with other systems, the study selects an analytical method to capture the interrelations between key factors from the macroscopic perspective.

Many experimental results reported in the literature show that the capacity drop is approximately 5 to 30 percent below from the theoretical capacity under the free-flow condition [100-103, 133-135]. Several investigations attribute such capacity drop to the patterns of slower merging vehicles that create local voids during the merging process and consequently reduce the traffic flow rate on the open lane(s) [136]. Some studies in the literature [90, 137-140] tried to develop analytical models to describe those additional capacity drop phenomenon for the ramp-merging area based on the bounded acceleration capabilities of vehicles, because the slow insertion of merging vehicles often creates a void in the traffic stream and thus reduces the available roadway capacity [96]. However, different from the on-ramp area, to describe the capacity reduction phenomenon on a lane-closure highway segment by an incident, one needs to deal with several complex interaction among traffic flows. For instance, the merging points of lane-changing vehicles are widely distributed along the upstream roadway segment of the lane-closure location. Thus, merging vehicles from different merging locations may have different impacts on the capacity drop. Moreover, since traffic conditions in the incident scene are mostly at the state of the congestion, merging vehicles, if executing the lane changes sufficiently distant from the lane-closure spot, may have enough time to catch up its created void in the traffic stream. Thus, those merging maneuvers may not contribute to the capacity drops at the lane-closure spot.

Therefore, using the analytical approaches, this study will describe those unique but complex relations between key traffic characteristic variables under the lane-closure conditions, and develop a set of convenient models for use in practice to estimate the additional capacity reduction.

7.2. Model development

Despite the fact that many factors can affect the available capacity at the highway lane-closure location, drivers' lane-changing and merging behaviors from the close lane onto the open lane is the most critical one [95]. Such merging behaviors may also vary with a vehicle's merging speed, distance to the blockage location, volume ratio between the open and close lanes, and the acceleration rate. Therefore, an analytical model for estimating the reduced roadway capacity should be capable of reflecting the inter-relations between those key factors associated with merging behaviors and the resulting discharging flow rate.

This chapter first presents a base model for the single-lane-open scenario, focusing on establishing the relations between the key contributing factors resulting in the capacity drop. A more generalized model for use in the multi-lane-open scenarios on highway segments experiencing one or multi-lane closure will be discussed in the next chapter.

7.2.1. Overview and assumptions

Figure 30 illustrates the relations between key contributing factors and their interactions. Table 23 lists all related parameters and variables used for the base model development.

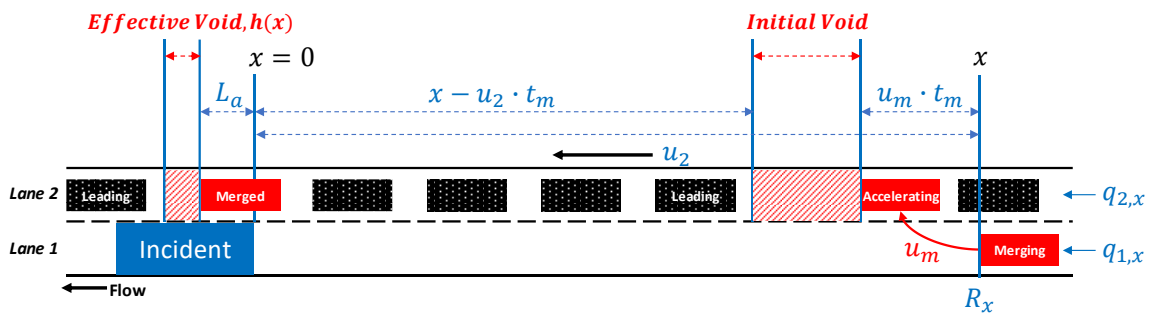


Figure 30. Illustration of a single-lane-open scenario and the key factors

Table 23. Notations for a capacity drop model

Variable	Definition
u_i	vehicle speed on Lane i (m/s)
u_m	merging speed of vehicles (m/s)
u_f	theoretical free-flow speed (m/s)
w	theoretical wave-flow speed (m/s)
k_j	theoretical jam density (veh/m)
x	merging location from the lane-closure point (meters)
t_m	merging time (sec) with mean \bar{t}_m and variance S_{tm}
a	acceleration rate (m/s ²) with mean \bar{a} and variance S_a
p_h	proportion of heavy vehicles
$q_{i,x}$	incoming flow rate on Lane i at x (veh/s)
$q_{i,i+1}$	merging flow rate form Lane i to $i+1$ (veh/s)
$R_{i,i+1,x}$	merging ratio from Lane i to $i+1$ at x
$L_{a,i}$	reference point to compute an effective void on Lane i (meters)
$L_{max,i}$	maximum effective merging location on Lane i (meters)
$L_{min,i}$	minimum effective merging location on Lane i (meters)
$P_{m,i}(x)$	probability of vanishing a void on Lane i where merging at x
$h_i(x)$	effective void on Lane i without interactions where merging at x (meters)
$h_i^l(x)$	effective void on Lane i with interactions where merging at x (meters)
C_i^e	effective capacity of Lane i (veh/s)
C_i^t	theoretical capacity of Lane i (veh/s)

All terminologies used hereafter in model description are defined in detail as follows:

- **Theoretical capacity (vphpl), C_i^t** : the maximum observed flow rate under the free-flow condition.
- **Effective capacity (vphpl), C_i^e** : the maximum outflow rate (i.e., available capacity) from the lane closure spot due to an incident and traffic conflicts.
- **Capacity drop (vphpl), $C_i^t - C_i^e$** : the difference between the theoretical capacity and effective capacity.

- **Merging ratio, $R_{i,i+1,x}$** : the number of merging vehicles divided by the number of vehicles on the target lane (i.e., $i + 1$) lies between the lane-closure spot (i.e., $x = 0$) and the merging location of vehicles (i.e., x).
 - Note that the merging ratio at any location in the single-lane-open scenario (or, mandatory lane changes) can be computed with the incoming flow rates from the open and closed lanes, because all vehicles on the closed lane must merge into the opened lane before reaching the lane-closure spot (i.e., $R_x = q_{1,x}/q_{2,x}$).
- **Void (meter)**: the part of space headway that is larger than the uninterrupted (without the merging-in vehicles) space headway in the traffic flows.
- **Initial void (meter)**: the void created initially between a leading vehicle and a merging vehicle in traffic stream due to a slow insertion of the merging vehicle.
- **Effective void (meter), $h_i(x)$** : the void between a leading vehicle and a merging vehicle, when the vehicle merging from a location, x , arrived at a reference point, L_a , associated with the lane-closure spot.
- **Reference point (meter), L_a** : a critical point measured from the lane-closure spot for computing an effective void (i.e., $u_{m,at\ ref.} \geq u_{open}$ ¹).
 - The logic is that a merging vehicle closest to the lane-closure spot will have the most significant impact on the resulting capacity drop. Hence, the reference point is defined as the point where a void created by a merging vehicle will reach its maximum.

¹ $u_{m,at\ ref.}$: speed of a merging vehicle at the reference point; u_{open} : speed of traffic flow on the open lane.

- **Minimum effective merging location (meter), L_{min} :** a minimum distance, from the lane-closure spot, needed for a merging vehicle to merge into the target lane, that is, all vehicles should complete the merge before reaching the lane-closure spot.
- **Maximum effective merging location (meter), L_{max} :** a maximum distance from the lane-closure spot, beyond which vehicles can have enough time to catch up with their created void. Hence, such merging maneuvers taking place ahead of L_{max} will not cause the capacity drop at the lane-closure spot.
- **Probability of vanishing a void, $P_{m,i}(x)$:** the probability that a void created by a merging vehicle at location, x , will be compressed to a normal space headway before reaching the lane-closure spot.

This study will first describe the core logic of the capacity drop due to a lane-closure under the following assumptions. As shown in Figure 30, traffic flows on the highway segment are assumed to be homogeneous (i.e., all passenger cars), and follow a triangular fundamental diagram with free-flow speed u_f , wave speed w , and jam density k_j [141, 142]. The initial speed of vehicles in each lane is denoted as u_i , and the merging speed of a vehicle from the close lane to merge into the open lane at location x is defined as u_m , with its merging time of t_m . This potential merging location, x , is assumed to uniformly distribute along the roadway segment after vehicles start to merge [143]; the merging speed u_m shall be slower than the traffic flow speed, u_2 , of the target lane [90]. All vehicles in the closed lane must merge into the open lane before reaching the lane-closure spot. Note that when a merging vehicle starts to merge into the target lane with a merging angle, its speed is always slower than the speeds of those vehicles in the target lane (i.e., $u_m < u_2$),

even if the merging vehicle does not reduce its speed during the merging process. Hence, such a merging process generally creates an excessive space headway (named ‘void’) unused by traffic flows in the open lane, and then accelerates to catch up this created void (i.e., additional space headway) before reaching the lane-closure area, as shown in Figure 31.

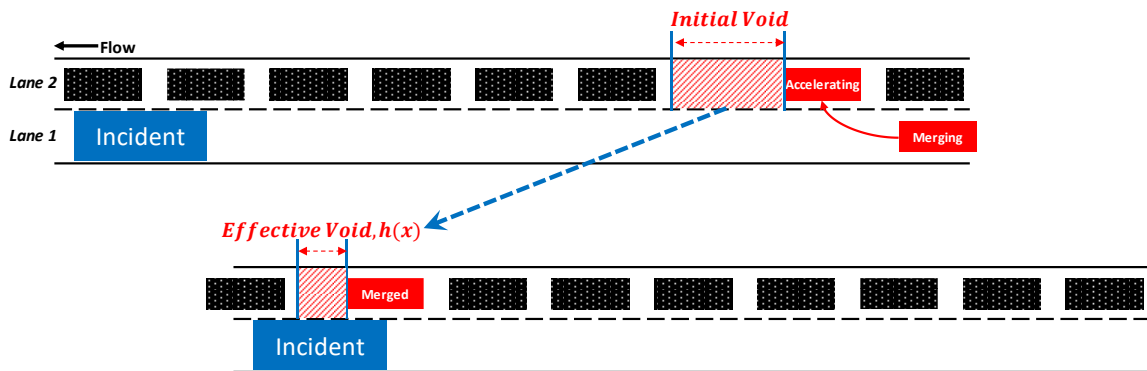


Figure 31. Illustration of a relation between the initial void and effective void

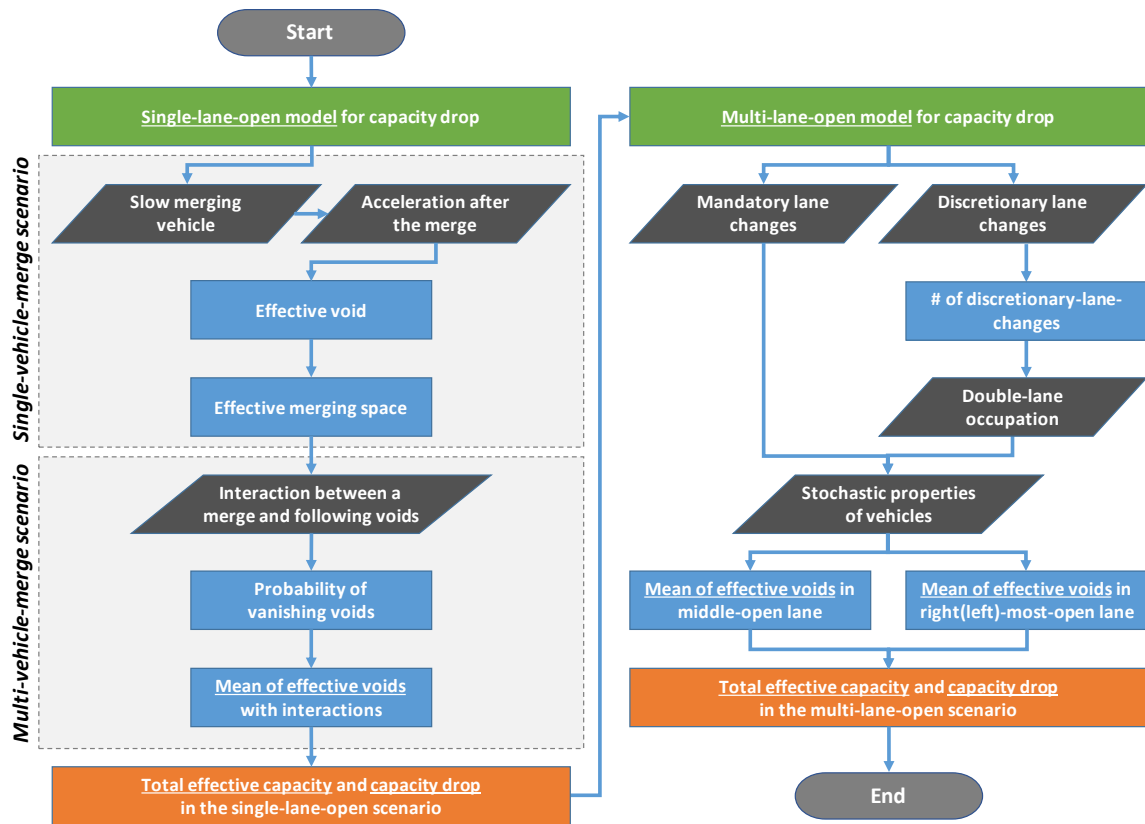


Figure 32. Framework for developing a capacity-drop model

As shown in Figure 32, this chapter and the following chapter will describe key ideas and concepts for the capacity drop by an incident and developed a single-lane-open-model and a multi-lane-open model for estimating the capacity drop as follows:

<Single-lane-open model>

- **Step-1:** set a single-vehicle-merge scenario; and compute the ‘effective void’ generated in the traffic stream due to merge-in vehicles. Such an effective void, the major cause for capacity drop, varies with the void created by a slow-merging vehicle and acceleration time (or, merging location from the lane-closure area) needed for a merging vehicle to catch its created void.

- **Step-2:** estimate an ‘effective merging space’, defined as the segment of a highway measured from the lane-closure spot within which a merging maneuver taking place will cause capacity drop.
- **Step-3:** define multi-vehicle-merge scenarios; and formulate the interactions between the void created by a merging vehicle and its preceding merging vehicles in the multi-vehicle-merging scenarios.
- **Step-4:** compute the probability that a void will be compressed to a normal headway due to vehicles merging into the open lane ahead of it, based its merging location.
- **Step-5:** apply the probability computed in Step-4 to the effective void computed in Step-1 to reflect such impacts due to multiple vehicle merges.
- **Step-6:** compute the mean of effective voids with the probability from Step-5 and the effective merging space estimated in Step-2.
- **Step-7:** estimate the total effective capacity and capacity drop in the single-lane-open scenario due to a traffic incident.

<Multi-lane-open model>

- **Step-1:** classify the merging maneuvers in a multi-lane-open scenario into mandatory and discretionary lane changes.
- **Step-2:** compute the number of discretionary-lane-changing vehicles caused by the speed differences between adjacent lanes.
- **Step-3:** formulate the double-lane occupied process by a merging vehicle, and the resulting impacts on the available roadway capacity with respect to their stochastic properties such as merging time and acceleration rate.

- **Step-4:** compute the mean of effective voids in the middle-open lanes and the right(left)-most-open lane by sequential adoption of the single-lane-open model for the mandatory lane changes, the discretionary-lane-changes from Step-2, and the key properties in Step-3.
- **Step-5:** estimate the total effective capacity and capacity drop in the multi-lane-open scenario due to a traffic incident.

7.2.2. Computation of effective voids

To capture the relations between the merging maneuvers and the resulting capacity drop, one can first assume that only one merging maneuver between the closed and open lanes will take place during each time interval. Since a merging vehicle typically creates a ‘void’ in the traffic stream due to its relative slow movements and bounded acceleration constraints during the merging process, such a void will in turn reduce the available space on the open lane and thus result in temporary flow restrictions. After completing the merge, such a merging vehicle will start to accelerate and compress its created void, and then catch up its leading vehicles. Therefore, one can estimate the actual unused space headways in the lane-closure spot and the resulting capacity drop by computing the ‘effective void’.

$$h(x) = \max \left[\underbrace{u_2 \cdot \frac{-u_m + \sqrt{u_m^2 + 2a(x - u_m \cdot t_m + L_a)}}{a}}_{\textcircled{1}} - \underbrace{(x - u_2 \cdot t_m + L_a)}_{\textcircled{2}}, \quad 0 \right] \quad (6)$$

By using the kinematic equations, one can derive the effective void under the above scenario with Equation (6). Note that the first term of Equation (6), $\textcircled{1}$, indicates moving

distance of the leading vehicle on the target lane, when the merging vehicle arrives at the reference point (see Figure 30). This moving distance (i.e., ① of Equation (6)) can be computed by multiplying the vehicle speed on the target lane (i.e., u_2) and the time for the merging vehicle to reach the reference point (i.e., L_a). This travel time can be derived from the kinematic equation (e.g., $s = v_0 \cdot t - 0.5 \cdot a \cdot t^2$).

The second term of Equation (6), ②, shows the distance of the leading vehicle from the reference point when the merging vehicle completes the merging maneuver. Therefore, the effective void (i.e., $h(x)$), the distance between the preceding and the merging vehicles at the reference point, can be computed by subtracting the moving distance (i.e., ① of Equation (6)) from the distance from the reference point (i.e., ② of Equation (6)).

Note that, in Equation (6), the reference point for computing the effective voids (i.e., L_a) can be derived when the speed of a merging vehicle (i.e., u_m) with acceleration (i.e., a) evolves to the same speed as the traffic flows on the open lane (i.e., u_2), as shown in Equation (7). Also, each lane's flow speed (i.e., u_i) can be computed from the relations of the fundamental diagram with Equation (8) [141, 142].

$$L_a = \frac{(u_2^2 - u_m^2)}{2a} \quad (7)$$

$$u_i = \frac{(w \cdot q_i)}{(w \cdot k_j - q_i)} \quad (8)$$

7.2.3. Effective merging space

From those equations and relations, one can estimate the effective merging space, based on the minimum and maximum effective merging locations, as shown in Figure 33.

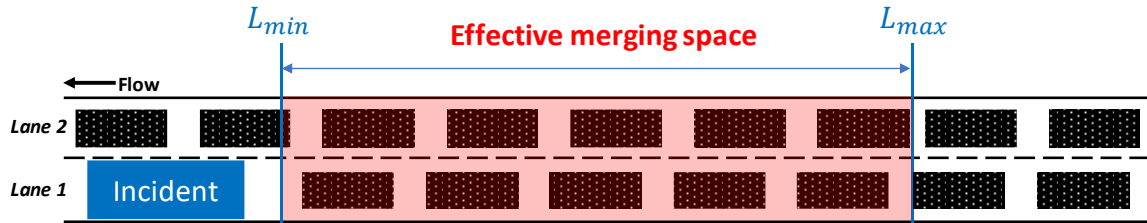


Figure 33. Illustration of effective merging boundaries

$$L_{min} = u_m \times t_m \quad (9)$$

$$h(L_{max}) = 0 \quad (10)$$

The ‘minimum effective merging location’ (i.e., L_{min}) is the minimum distance from the lane-closure spot needed for a merging vehicle to merge into the target lane, indicating that no merging maneuver shall take place beyond this location. This minimum effective merging location can be derived with the merging speed, u_m , and the merging duration, t_m , as shown in Equation (9).

The ‘maximum effective merging location’ (i.e., L_{max}) implies that a merging vehicle can have enough time to catch up with its created void if it merges into the open lane before the point, L_{max} . This maximum effective merging location can be computed by setting an effective void to be zero with Equation (10).

Finally, one can define the effective merging space as shown in Figure 33, where merging maneuvers taking place within such bounds will cause additional reduction in the roadway capacity.

7.2.4. Computation of effective voids under multiple vehicle merges

This section extends the formulations from a single merge to multiple merges on a lane-closure highway segment. When a vehicle starts to merge into the open lane, vehicles behind its merge point in the open lane will be forced to slow down. Thus, the void created by other early merging vehicles but behind the subject merging vehicle will be compressed to a smaller size or even to the normal headway of the traffic flows, as shown in Figure 34. Hence, the interactions between a subject merging vehicle and a void already created on the open lane by the earlier merging vehicles needs to be formulated in the multiple-merging scenarios.

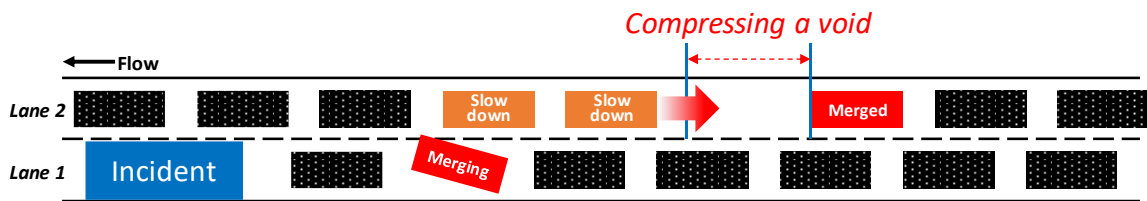


Figure 34. Illustration of interaction between an existing void and a vehicle merging ahead of its location

The probability of compressing a void to the normal headway, is a function of the merging ratio over the roadway segment. Because it depends on whether such a void will be compressed by another merging vehicle before it reaching the lane-closure area. Therefore, the probability of vanishing a void at the maximum effective merging location, L_{max} , is the same as the merging ratio, $R_{L_{max}}$, because a void from this location should meet another merging vehicle from the close lane before reaching the lane-closure area. Likewise, the probability of vanishing a void at the minimum effective merging location, L_{min} , is zero because no merging maneuver will take place after this minimum effective merging location.

Note that, if the merging locations of vehicles on the close lane follow a certain probability density function $f_m(x)$, along the effective merging boundaries between L_{min} and L_{max} , then their created voids shall follow the cumulative distribution function, $F_m(x)$. In addition, if the merging locations along the merging boundaries are assumed to be distributed uniformly, the probability of vanishing a void is given by Equation (11).

$$P_m(x) = F_m(x) = \int_{L_{min}}^x f_m(x) dx = \frac{R_{L_{max}}}{L_{max} - L_{min}} \cdot (x - L_{min}) \quad (11)$$

Therefore, an effective void, considering the interactions between a void and vehicles merging ahead of it (i.e., $h^I(x)$), can be computed by multiplying the effective void without the interactions (i.e., $h(x)$) and the probability of not vanishing a void (i.e., $1 - P_m(x)$), as shown in Equation (12).

$$h^I(x) = h(x) \times (1 - P_m(x)) \quad (12)$$

Lastly, the mean of effective voids can be calculated as shown in Equation (13) since the probability density function $f_m(x)$ is the uniform distribution.

$$E[h^I(x)] = \frac{\int_{L_{min}}^{L_{max}} h^I(x) \cdot f_m(x) dx}{\int_{L_{min}}^{L_{max}} f_m(x) dx} = \frac{\int_{L_{min}}^{L_{max}} h^I(x) dx}{L_{max} - L_{min}} \quad (13)$$

7.2.5. The effective capacity for single-lane-open scenarios

The effective capacity can be computed by measuring the size of such effective voids created by those merging vehicles, and the percentage of roadway capacity is unused due to those created voids. Therefore, one can formulate the effective capacity with Equation (14).

$$\begin{aligned}
 C^e &= C^t \times \frac{(q_{2,L_{max}} + q_{1,L_{max}}) \cdot \frac{1}{k_2}}{\underbrace{(q_{2,L_{max}} + q_{1,L_{max}}) \cdot \frac{1}{k_2}}_{\textcircled{1}} + \underbrace{q_{1,L_{max}} \cdot E[h^l(x)]}_{\textcircled{2}}} \\
 &= C^t \times \frac{(1 + R_{L_{max}})}{(1 + R_{L_{max}}) + R_{L_{max}} \cdot k_2 \cdot E[h^l(x)]}
 \end{aligned} \tag{14}$$

The first term of Equation (14), $\textcircled{1}$, is the same as the numerator of the equation, indicating the total occupied spaces by the incoming vehicles. It can be computed by multiplying the total number of incoming vehicles (i.e., $q_{1,L_{max}} + q_{2,L_{max}}$) and their average space headway (i.e., $1/k_2$). The second term of Equation (14), $\textcircled{2}$, indicates the total unused spaces (or, total effective voids) due to the merging vehicles that can be computed by multiplying the total number of merging vehicles (i.e., $q_{1,L_{max}}$) and the mean of effective voids of merging vehicles (i.e., $E[h^l(x)]$). Thus, the effective capacity, C^e , can be computed by multiplying the theoretical capacity (i.e., C^t) and the fraction of the maximum theoretical capacity unused by the all effective voids of the merging vehicles. Thus, Equation (14) can also be described with the merging ratio, $R_{L_{max}}$, from its definition in the single-lane-open scenario.

7.3. Numerical analysis of the model properties

7.3.1. Results from a numerical example

To analyze the mathematical properties of the proposed models, this section presents some numerical results using the case and key parameters reported in the literature (Leclercq & Laval [138] and Leclercq et al. [136]), as shown in Table 24.

Table 24. Variables and parameters for the numerical analyses

Variable	Parameter
Theoretical capacity C^t :	2400vphpl
Free-flow speed u_f :	115km/h
Jam density k_j :	145vpkpl
Merging ratio of a mandatory lane changing $R_{3,Lmax,3}$:	0.5
Merging speed u_m :	10km/h slower than u_i
Acceleration rate of vehicles a :	$2m/s^2$
Merging time t_m :	5sec

From this reference scenario, the estimated effective capacity with the developed model is about 1983vphpl, a reduction of about 17.4% from its capacity of 2400vphpl. The estimated capacity reduction with the developed capacity-drop model is within the same range of 5% and 30% as those reported in the literature [100-103, 133-135].

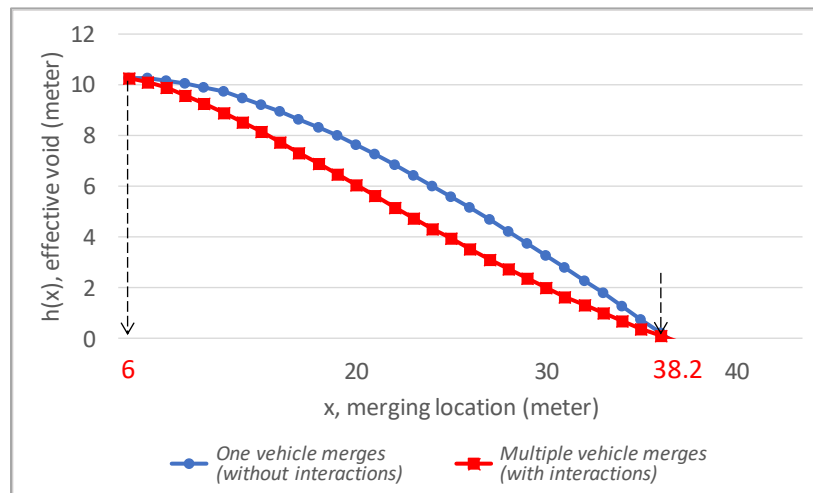


Figure 35. Effective voids vs. merging location with and without interactions between a void and vehicles merging ahead of its location

Figure 35 shows the effective voids without/with interactions between a void and vehicles merging ahead of its location. The numerical results reveal the following relations:

- The effective void decreases along with the merging location of lane-changing vehicles. Thus, the merging maneuvers ahead of the maximum effective merging location does not have any impact on the available-lane capacity due to the sufficient time duration for them to catch up the created void.
- The effective voids, considering the interactions among the preceding or subsequently created voids, is slightly less than the effective voids without the interactions. Because the amount of the created effective voids can be reduced when another vehicle is merging in front of them. This finding is in consistent with the previous studies [136, 139].
- The effective merging spaces remain unchanged regardless of those interactions. The minimum and maximum effective merging locations are 6 and 38.2 meters, respectively, under both merging scenarios.

Those vital properties, quite informative for design of dynamic merge control-strategies, have not ever been analyzed and reported yet in the literature.

7.3.2. Sensitivity analysis

Further investigation of the relations among the key contributing factors and the capacity drops has been conducted via extensive sensitivity analyses in Figure 36 with respect to the merging speed u_m , merging time t_m , acceleration rate a , and merging ratio R_{Lmax} on the capacity drop.

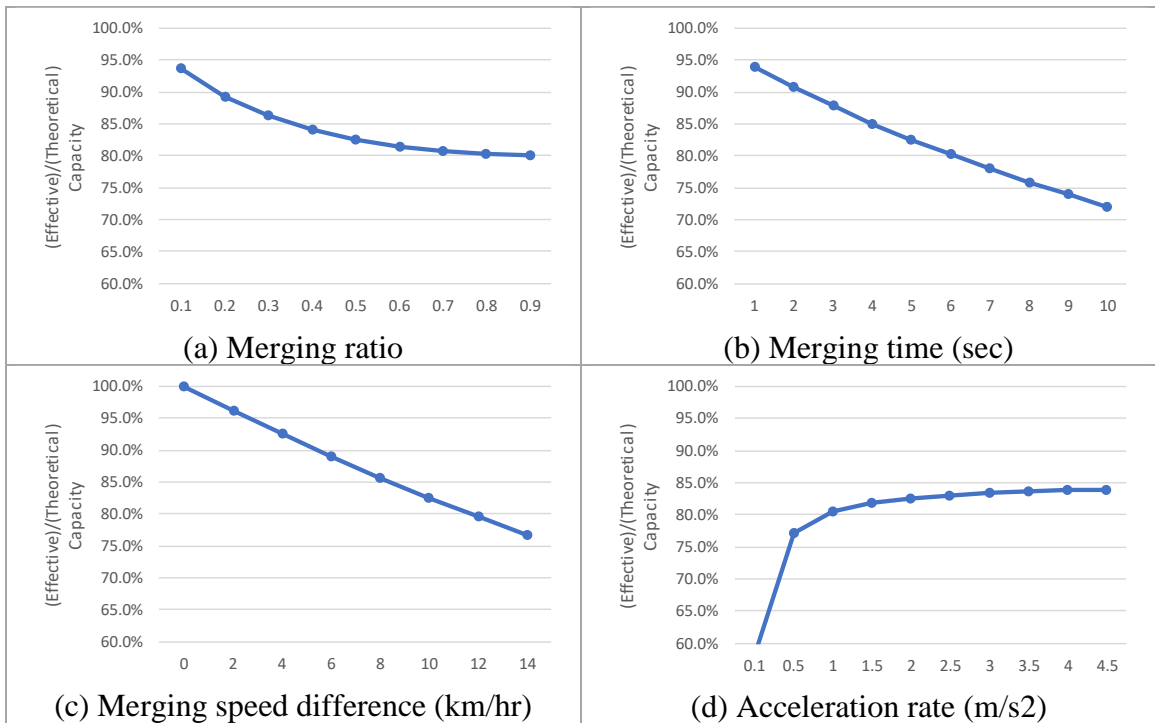


Figure 36. Sensitivity analysis of four variables on the capacity drop

Figure 36 (a) shows that the effective capacity decreases with an increase in the merging ratio. The rates of such changes in the effective capacity are less than 5%, when the merging ratios are larger than 0.4 (i.e., around 85% and 80%). It means that a large portion of the merging vehicles will not significantly cause the capacity drop under the scenario of high merging ratios (i.e., the merging ratio is larger than 0.4), because most of those voids created within the effective merging boundaries will be compressed and

eventually vanished before those vehicles reach the lane-closure spot. Only the vehicles that merge near the lane-closure spot can significantly contribute to the capacity drop.

From Figure 36 (b), it is noticeable that the effective capacity decreases when the merging time needed to complete a merge increases, because such merging vehicles will create larger voids during the merging process. The effective capacity as shown in Figure 36 (c) decreases with the merging speed difference of vehicles since the merging vehicles may take more time to catch up with the voids in front of them.

Figure 36 (d) shows that the effective capacity increases slightly with the acceleration rate of the merging vehicles. For accelerations above 1 m/s^2 , the acceleration rate does not significantly play any role in causing the capacity drop. This implies that such merging vehicles can easily catch up with the voids in front of them when the acceleration rate of the merging vehicles is at least 1 m/s^2 . Note that the speed difference between the merging vehicle and the vehicle on the target lane is only about 10 km/h in the example, so a higher acceleration rate cannot make a considerable difference with respect to the time duration needed to catch up with the voids. This result also indicates that heavy vehicles with a lower acceleration rate will have significantly impacts on the capacity drop.

7.4. Interim findings

This chapter has presented an analytical model for estimating the capacity drops of single-lane-open scenarios with homogeneous flows. By using the parameters that significantly affect the merging behaviors at the lane-closure highway segment (such as vehicle speed, acceleration rate, merging time, and merging ratio), this study has calculated effective voids and effective merging boundaries which in turn are used as the input for the function derived to estimate the effective roadway capacity under lane closures.

The research findings from this chapter are summarized below:

- The developed single-lane-open model for estimating the capacity drop due to an incident can be used to describe the relations between the capacity drop and its contributing factors. The developed model shows that:
 - Each merging maneuver depending on its merging location has a different impact on the capacity drop.
 - A merging maneuver closest to the lane-closure spot has the most significant impact on the capacity drop, and a merging maneuver taking place ahead of the maximum effective merging location will not cause any drop in the open-lane's capacity.
- The effective capacity is correlated inversely with the merging time and the merging ratio, but increases the merging speed and acceleration rate.
- The effective capacity is quite sensitive to the low acceleration rate, reflecting that heavy vehicles with a lower acceleration rate can have a huge impact on the capacity drop.

Chapter 8. A generalized model for a lane-closure highway segment: multi-lane-open scenarios

8.1. Model development

8.1.1. Overview and assumptions

Figure 37 shows a typical lane-closure highway segment that has multiple open lanes to accommodate the incoming lane-changing vehicles, where there are two types of vehicles, such as passenger cars and heavy vehicles, with different acceleration and deceleration rates.

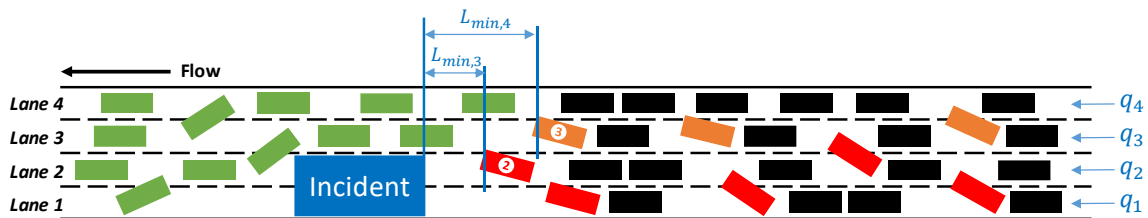


Figure 37. Illustration of merging behaviors in the multi-lane-open scenario

The capacity drop under such multi-lane-open scenarios can be attributed to two types of lane-changing behaviors: mandatory and discretionary lane changes. First, all vehicles on the closed lanes (i.e., Lanes 1 and 2) must move to the open lanes (i.e., Lanes 3 and 4) due to the blockage. Such lane change maneuvers are denoted as ‘mandatory lane changes’, and those vehicles are colored in red in Figure 37. Also, some vehicles on the open lane next to the blocked lane (i.e., Lane 3) may choose to move to the neighboring lane (i.e., Lane 4) to avoid a speed reduction due to interruptions of incoming flows from the blocked lane. Those lane-changing maneuvers are defined as ‘discretionary lane changes’, and those vehicles are colored in orange in Figure 37. As both types of lane-changing maneuvers spread the queue laterally across the highway segment, they will inevitably cause additional capacity drop [104].

After passing the lane closure spot (i.e., green colored vehicle in Figure 37), vehicles, slowing down due to the queues, start to discharge and speed up since all lanes are available again. Thus, traffic flows in the downstream segment of a lane-blockage spot are generally moving under a free-flowing condition [104].

8.1.2. Effective merging boundaries for each lane

Using the same logic as in the single-lane-open scenario, one can also compute the minimum distance for vehicles to complete the merges in each lane under the multi-lane-open scenarios. As shown in Figure 37, Vehicle-2 in Lane 2 must start to merge to Lane 3 before reaching the minimum effective merging location, $L_{min,3}$, the same concept used in the single-lane-open analysis. However, Vehicle-3 in Lane 3 has the option to merge to the right lane to avoid the slow moving-in vehicles from Lane 2. Thus, such vehicles in Lane 3 also need to complete the merges before reaching the critical location, $L_{min,4}$. Note that to avoid potential conflicts with the merging-in vehicles, those vehicles in the right lane for non-mandatory lane changes typically are constrained to complete the maneuvers more distant ahead of the blockage spot than those vehicles in the left lane for mandatory lane changes. For instance, $L_{min,4}$ is always further away from the blockage spot than $L_{min,3}$ (e.g., $L_{min,3} < L_{min,4}$). Lastly, those minimum and maximum effective merging locations for each lane can be derived from the equations of the single-lane-open model.

Note that since the focus is to estimate the additional capacity drop due to the interference of merging vehicles to the open-lane flows, one needs not to consider the lane-changing maneuvers between those blocked lanes.

8.1.3. Double-lane occupied process by one merging vehicle

With respect to the lane-changing process on the open lanes, it is clear that vehicles (as shown in Figure 38) may concurrently occupy two lanes, and thus creates voids on those two lanes, resulting in a capacity reduction to both lanes.

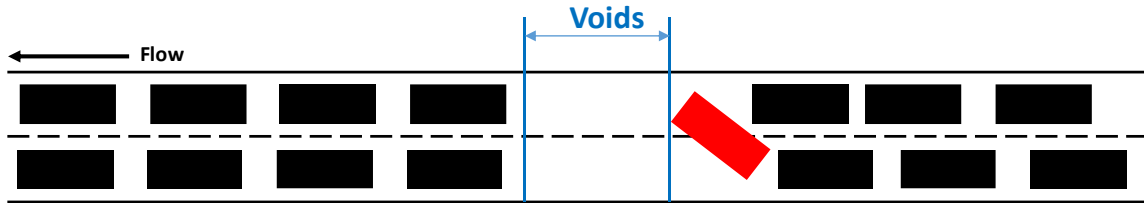


Figure 38. Illustration of a double-lane occupied process on two open lane by one merging vehicle

8.1.4. Stochastic properties

Note that each lane under lane-closure scenarios is most likely to be under the congested condition. Hence, the merging speeds, u_m , and vehicle speeds on the same blockage lane, u_i , can both be viewed as to follow a homogeneous distribution [136], but not for merging times t_m and acceleration rates a which actually vary with different types of vehicles. Therefore, this study denotes the heterogenous merging time and acceleration rate as $t_m(\bar{t}_m, s_{tm} > 0)$ and $a(\bar{a}, s_a > 0)$, respectively. By denoting the percentage of heavy vehicles as, p_h , one can also compute those merging times and acceleration rates with Equations (15) and (16), where $t_{m,p}$ and $t_{m,h}$ denote the merging times for passenger cars and heavy vehicles, respectively; a_p and a_h are defined as the acceleration rates for passenger cars and heavy vehicles, respectively.

$$t_m = (1 - p_h) \cdot t_{m,p} + p_h \cdot t_{m,h} \quad (15)$$

$$a = (1 - p_h) \cdot a_p + p_h \cdot a_h \quad (16)$$

Hence, considering the heterogeneous nature of traffic flows, one can update the effective void derived from Equation (6) for homogeneous flows to Equation (17) by using the multivariate generalization of the delta method, assuming that the covariance between the acceleration and merging time is zero [144].

$$h(x)_{s_a > 0, s_{tm} > 0} = \max \left[h(x, \bar{a}, \bar{t}_m) + \frac{1}{2} s_a^2 \frac{\partial^2 h}{\partial a^2} + \frac{1}{2} s_{tm}^2 \frac{\partial^2 h}{\partial t_m^2}, 0 \right] \quad (17)$$

8.1.5. The effective capacity for multi-lane-open scenarios

Notably, the capacity drop in the multi-lane-open scenarios can be calculated with the same notion as to measuring the outgoing flows from each open lane under both the mandatory and discretionary lane-changing interferences.

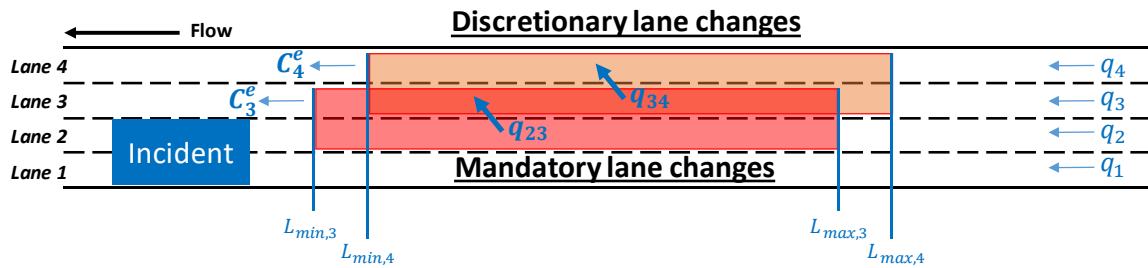


Figure 39. Discretionary and mandatory lane changes and key variables

Figure 39 shows the example of a highway segment where its two out of four lanes are closed due to an incident, and its two types of lane-changing maneuvers may cause additional capacity drop, i.e., q_{23} from Lane 2 to Lane 3 and q_{34} from Lane 3 to 4. Therefore, the total effective capacity C_{total}^e can be calculated by the effective capacity for each open lane C_i^e with Equation (18).

$$C_{total}^e = \sum_{i \in open} C_i^e = C_3^e + C_4^e \quad (18)$$

8.1.5.1. Effective capacity of the right(left)-most lane

Since for either the leftmost or rightmost open lane, it will only receive incoming merging flows (see Figure 40), one can thus compute its effective capacity (C_4^e) by accounting for the impacts by merge-in vehicles. The capacity drop process on such lanes is similar to the process under the single-lane-open scenarios, so its effective capacity can be approximated with Equation (19).

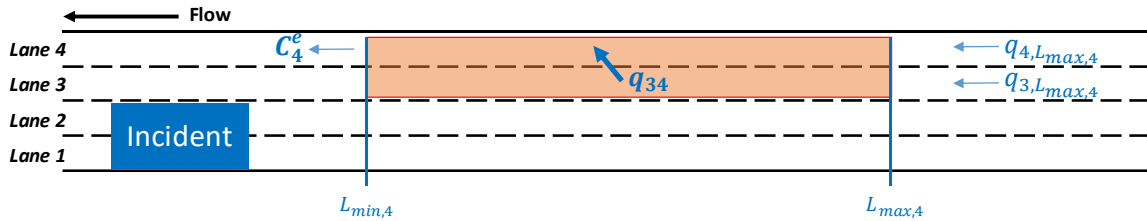


Figure 40. Key variables to estimate the effective capacity of the rightmost lane

$$C_4^e = C_4^t \times \frac{(q_{4,L_{max,4}} + q_{34}) \cdot \frac{1}{k_4}}{(q_{4,L_{max,4}} + q_{34}) \cdot \frac{1}{k_4} + q_{34} \cdot E[h_{4,s_{a,stm}}^I(x)]} \quad (19)$$

The number of vehicles that are likely to execute discretionary lane changes can be approximated with the macroscopic lane-changing model by Laval & Leclercq [145], as shown in Equation (20). The logic behind this macroscopic lane-changing model is that the percentage of the discretionary lane changes will be the function of the speed difference between two adjacent lanes [145].

$$q_{34} = C_4^e \cdot \frac{\max(u_4 - u_3, 0)}{u_f^2 \cdot t_m} \cdot L_{max,4} \quad (20)$$

8.1.5.2. Effective capacity of the middle lane

In any middle lane among the opened lanes, the following two types of merging behaviors can concurrently cause the additional capacity drop: i.e., incoming vehicles merging into the open lane and outgoing vehicles to the neighboring lane to avoid a speed reduction. Note that outgoing merging vehicles can still affect the capacity drop for the open lane because such merging vehicles will occupy two lanes over some time duration. Figure 41 shows lane-changing flows from Lanes 2 to 3, q_{23} , and from Lanes 3 to 4, q_{34} , and both may affect the capacity of Lane 3. Therefore, the effective capacity of any middle open lane can be calculated by accounting for those two merging maneuvers as shown in Equation (21).

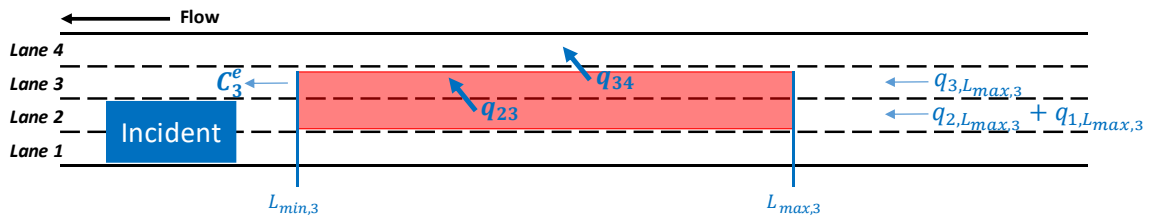


Figure 41. Key variables to estimate the effective capacity of the middle lane

$$C_3^e = C_3^t \times \frac{(q_{3,Lmax,3} + q_{23} - q_{34}) \cdot \frac{1}{k_3}}{(q_{3,Lmax,3} + q_{23} - q_{34}) \cdot \frac{1}{k_3} + (q_{23} + q_{34}) \cdot E[h_{3,s_a,stm}^I(x)]} \quad (21)$$

Finally, using Equation (19) for the leftmost or rightmost lane and Equation (21) for each middle lane, one can use Equation (18) to compute the total effective capacity for all open lanes under the multi-lane-open highway blockage scenarios. Besides, from Equations (19) and (21), one can notice that the capacity reduction (total voids) in Lane 3 is always larger than Lane 4 ($C_4^e > C_3^e$), because the merging vehicles from Lane 3 to Lane 4 ($q_{3,4}$) also affect the capacity reduction in Lane 3. Such properties revealed from the analytical relations are consistent with the findings reported in the literature (Munjal et al. [146] and Bertini & Leal [105]).

8.2. Numerical analysis of the model properties

8.2.1. Results from numerical examples with different assumptions

To investigate the properties of the developed multi-lane-open model for capacity reduction with heterogeneity, the study has adopted the scenario reported in the literature (Leclercq & Laval [138] and Leclercq et al. [136]), as shown in Table 25.

Table 25. Variables and parameters for the numerical analyses

Variable	Parameter
Theoretical capacity C^t :	2400vphpl
Free-flow speed u_f :	115km/h
Jam density k_j :	145vpkpl
Merging ratio of a mandatory lane changing $R_{3,L_{max,3}}$:	0.5
Merging speed u_m :	10km/h slower than u_i
Percentage of heavy vehicles p_h :	15%
Acceleration rate of passenger cars $a_p(\bar{a}_p, sa_p)$:	(2m/s ² , 1)
Acceleration rate of heavy vehicles $a_h(\bar{a}_h, sa_h)$:	(1m/s ² , 0.5)
Merging time for passenger cars $t_p(\bar{t}_p, st_p)$:	(5sec, 2.5),
Merging time for heavy vehicles $t_h(\bar{t}_h, st_h)$:	(8sec, 4)

Table 26. Effective capacities estimated by the proposed model under different conditions

Conditions	No heavy vehicles Homogeneity	Heavy vehicles Homogeneity	Heavy vehicles Heterogeneity
	$p_h = 0\%$ $sa_p, sa_h, st_p, st_h=0$	$p_h = 15\%$ $sa_p, sa_h, st_p, st_h=0$	$p_h = 15\%$ $sa_p, sa_h, st_p, st_h>0$
C_4^e (in Figure 39)	2161vphpl (-10.0%)	2142vphpl (-10.7%)	2113vphpl (-11.9%)
C_3^e (in Figure 39)	1830vphpl (-23.8%)	1792vphpl (-25.3%)	1764vphpl (-26.5%)
Total effective capacity	3990vph (-16.9%)	3934vph (-18.0%)	3877vph (-19.2%)

Table 26 shows the resulting effective capacities for each lane, computed with the developed model, under different assumptions. First, the effective capacities of Lane 3 (i.e.,

middle-open lane) under all conditions are about 15% less than those of Lane 4 (i.e., right-most-open lane). Because the discretionary lane changes from Lane 3 to Lane 4 will affect the effective capacities of those two lanes concurrently. Such results are consistent with those revealed in the literature (Cassidy & Rudjanakanoknad [104] and Chung et al. [103]). Secondly, the total capacity drop increases about 16.9% to 19.2% when considering heavy vehicles and their resulting traffic flow properties. This is due to relatively low acceleration rates of the heavy vehicles and the additional interactions in the heterogeneous traffic flows often create larger voids during the merging process. Lastly, the total capacity drop, estimated with the developed model in the two-out-of-four-lanes scenario under different conditions, falls in the range of 16.9% to 19.2% less than the theoretical capacity (i.e., 2400vphpl), at the same level as those results estimated in other studies [100-103, 133-135].

8.2.2. Sensitivity analysis

To investigate how will the heterogeneous traffic flow properties affect the developed model's estimated capacity drop, this section presents the results of sensitivity analyses with respect to the percentages of heavy vehicles, standard deviation of acceleration rates, and standard deviation of merging times, as shown in Figure 42.

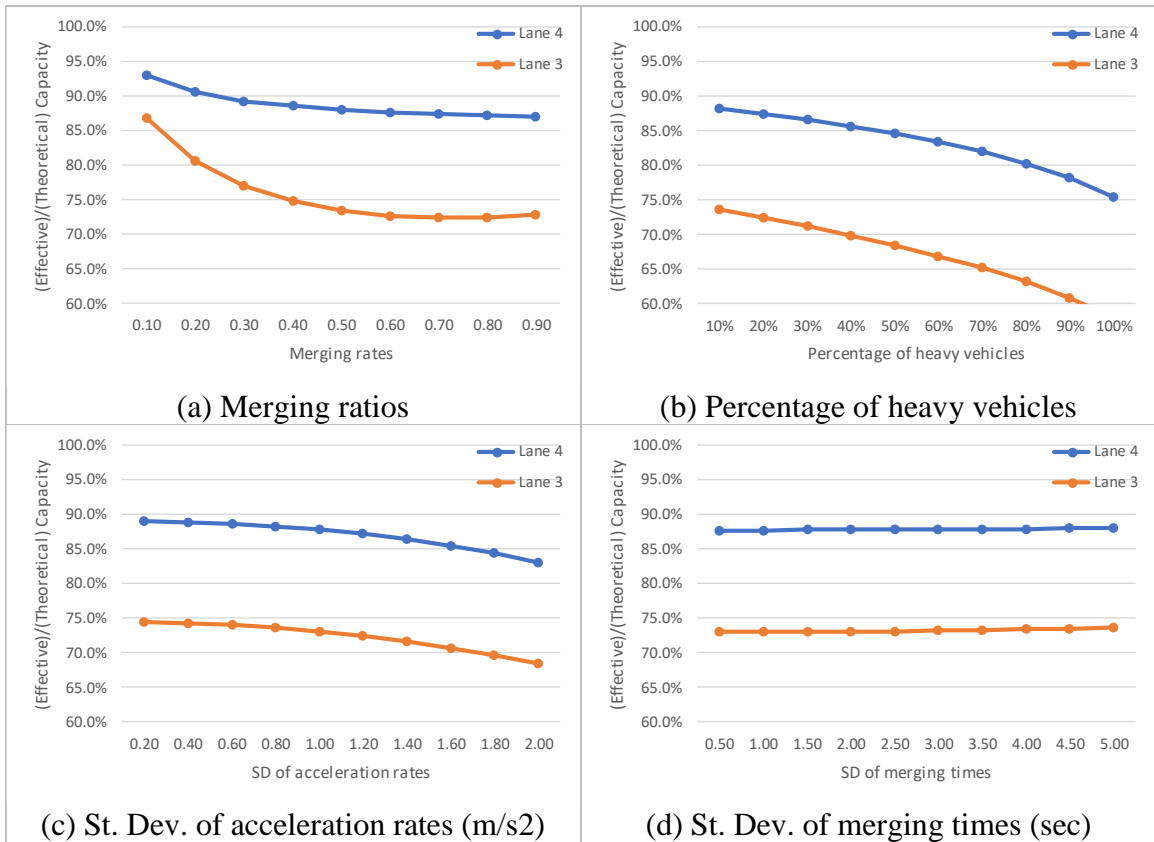


Figure 42. Sensitivity analysis of different parameters on the capacity drop

First of all, as shown in Figure 42 (a), the effective capacity exhibits a decreasing trend with respect to the increasing merging ratios. Notably, Lane 3, the closest one to the closed lane, is affected more by the changes in the merging ratio. However, the effective capacities of both lanes maintain at the stable level when the merging ratio is over 0.5. Such results

are similar to those numerical results for the single-lane-open model reported in the previous chapter. Secondly, the effective capacity significantly decreases as the percentage of heavy vehicles increases as shown in Figure 42 (b). This is due likely to the fact that heavy vehicles with their excessive weights and lengths often take more time to complete the merges. Such impacts on the capacity reduction are similar to the relationship between acceleration rate and effective capacity for the single-lane-open model. Thirdly, the effective capacity, as shown Figure 42 (c), slightly decreases with an increase in the standard deviation of the acceleration rate, indicating that discrepancies of acceleration rates in the traffic flows can cause additional drop in the effective capacity. Lastly, the results from Figure 42 (d) seems to show that the standard deviation of the merging times does not have significant impacts on the capacity reduction.

8.3. Experimental Validation

One major challenge in evaluating the capacity-reduction model during an incident scenario is the data availability and quality. This is due noticeably to the fact that an incident is not predictable, so one cannot plan a field survey of all essential data in advance. Therefore, this study utilizes RITIS and its CCTVs to collect the required data and assesses the developed capacity-reduction model's reliability during an incident.

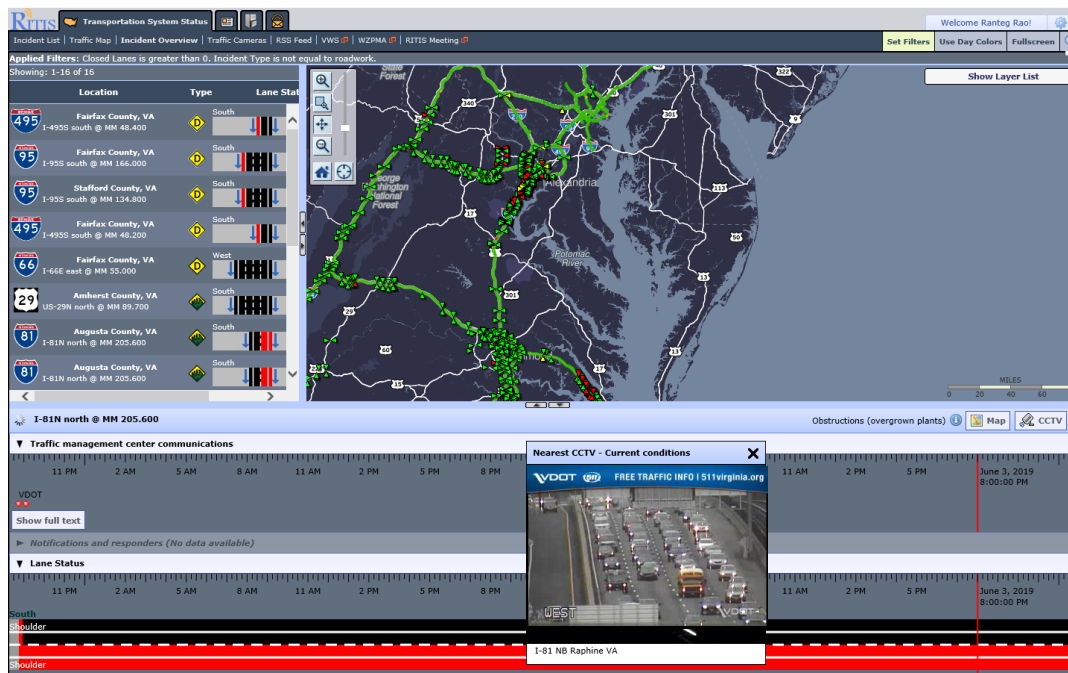


Figure 43. Examples for RITIS for incident managements

As shown in Figure 43, the RITIS contains a list of incidents and associated data, including the location, timeline, lane status, speed profile, activities of response units, CCTV, etc. Furthermore, the data in RITIS also includes the information of outflows from the lane closure and merging times by vehicle types observed from those CCTVs. Also, detectors near each identified incident can provide the traffic conditions and associated information during the incident clearance period for parameter calibration and assessment of the proposed model.

8.3.1. Incident information from RITIS

As shown in Table 27, three incidents and all associated data were collected from the RITIS and computed from the CCTV, and the results are shown in Table 28 and Table 29.

Table 27. List of incidents

Site	Date	Location	Lane Status	Detail
1	April 8, 2019 17:18~18:02	I-95 South, after Exit 73	1 of 3 lanes blocked (left)	CPD, 2 passenger cars, 1 CHART, 1 POLICE, 2 TOWs
2	April 5, 2019 15:52~17:00	I-64 West, after Exit 278	1 of 3 lanes blocked (right)	Collision (no info. for injury), 2 passenger cars, 3 CHARTs, 2 POLICE, 2 FIREs, 1 EMS, 1 TOW
3	May 7, 2019 14:24~15:01	I-64 West, after Exit 255	2 of 3 lanes blocked (right)	Vehicle on Fire, 1 passenger car, 2 CHARTs, 2 POLICE, 1 FIREs, 1 TOW

Table 28. Outflows from the lane closure by vehicle type

Site		Passenger car	Heavy vehicle	Total
1	Flow rate (vphpl)	1420	68×2*	1556
	Percentage Rate	91.25%	8.75%	100%
2	Flow rate (vphpl)	1435	49×2*	1534
	Percentage Rate	93.55%	6.45%	100%
3	Flow rate (vphpl)	1233	114×2*	1462
	Percentage Rate	84.34%	15.66%	100%

* the Passenger Car Equivalent (PCE) = 2.0 [147]

Table 29. Average merging times by vehicle type

Site		Passenger car	Heavy vehicle
1	Merging time (sec)	3.90	9.75
	Standard deviation	0.84	2.23
2	Merging time (sec)	3.95	9.86
	Standard deviation	0.86	2.25
3	Merging time (sec)	8.00	11.79
	Standard deviation	1.53	3.12

This study only considers two types of vehicles: passenger cars and heavy vehicles. Based on the FHWA’s vehicle classification scheme [148], passenger cars range from four-tire single unit vehicles to smaller vehicles (i.e., Class 1 to 3), and heavy vehicles include both buses and bigger vehicles (i.e., Class 4 to 13). No vehicles with six or more axle single-trailer (Class 10 to 13) were observed in the collected field data. In addition, the passenger car equivalent (PCE) of the heavy vehicles is set to 2.0 to account for their impacts, based on the 2016 HCM [147, 149]. Lastly, the merging speed of vehicles in a blocked lane is estimated with Equation (22), using the computed average merging time, lane width, and steering angle when making lane changes. Note that the standard lane width for the U.S. interstate highway system is 12-ft [150], and the mean steering angle for lane change is set to be 8.11 degree [151].

$$Merging\ speed = \frac{Lane\ width}{Merging\ time \times \tan(Merging\ angle)} \quad (22)$$

8.3.2. Detector data and estimated traffic conditions from RITIS

To estimate the fundamental diagram of traffic flows at the lane closure area, traffic volume and occupancy information, collected from the nearest detectors, are used to calibrate the q-k relation as shown in Table 30. The collected traffic conditions are divided into the free-flow conditions and congested conditions, based on the highest observed flows. Finally, one can estimate the observed maximum flow rate (i.e., the theoretical capacity), free-flow speed, and jam density of each location as shown in Table 30 from the regression results in Figure 44.

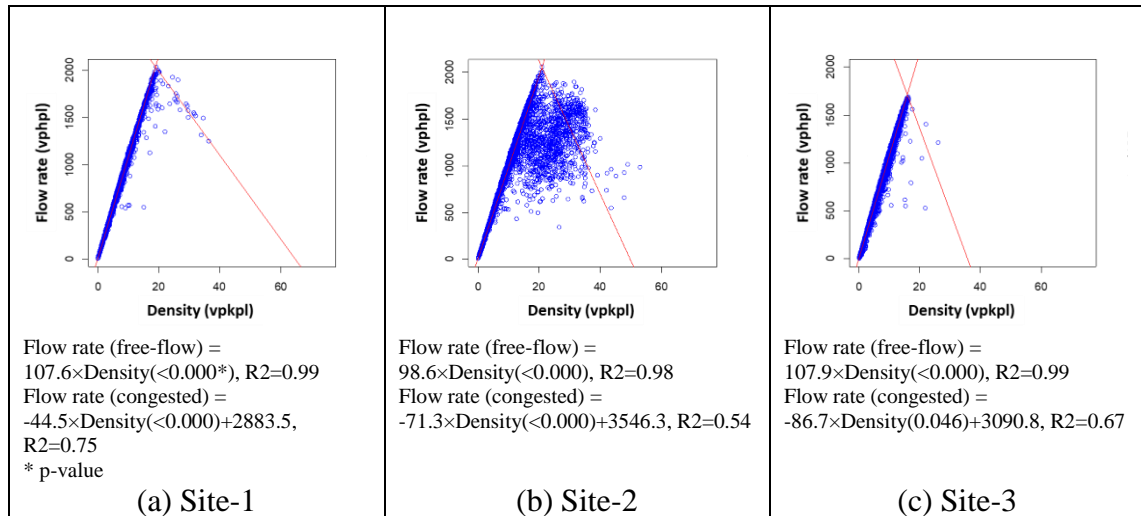


Figure 44. Scatter plots and regression results to estimate a fundamental diagram for each site

Table 30. Estimated traffic conditions of each site using the nearest detectors

site	date	Observed maximum flow rate (vphpl)	Free-flow speed (km/h)	Jam density (vpkpl)
1	Feb.~Jun., 2017*	2034.6	107.6	64.7
2	Feb.~Jun., 2018	2058.6	98.6	49.7
3	Jan.~Dec., 2018**	1694.3	107.9	35.6

* Detectors were deactivated during the year 2018

** Include more datapoints due to the lack of data in congested conditions

8.3.3. Comparison results

To assess the developed model, Table 31 shows the comparison between the field-observed and the model-produced results. Note that the acceleration rates of vehicles and their variances are unknown in the field data, thus parameters from 2016 HCM [147] were used as the inputs of the developed capacity-reduction model. The merging ratio was also assumed as 0.5. Those unknown parameters will be further discussed in the ensuing section.

Table 31. Comparison of the observations and model estimates

Site	Observed maximum flow rate (theoretical capacity, vphpl)	Under the lane-closure condition by an incident		
		Observed flow rate (vphpl)	Estimated flow rate (effective capacity from the model, vphpl)	Difference (observation-estimation)
1	2035 (100%)	1556 (76.5%)	1580 (77.6%)	- 1.2%
2	2059 (100%)	1534 (74.5%)	1522 (73.9%)	0.6%
3	1694 (100%)	1462 (86.3%)	1533 (90.5%)	- 4.2%

As shown in Table 31, the estimates from the developed model is very close to the observed capacity-reduction from the field data. All differences between the estimated and observed results are less than 5 percent. For example, from Site-1, the observed maximum flow rate from the detectors was 2,035 vphpl and the observed flow rate from the RITIS when the incident occurred was 1,556 vphpl, which is about 76.5% of the observed maximum flow rate. The estimated flow rate with the developed model is 1,580 vphpl, around 77.6% of the observed maximum flow rate, and about 1.2% deviated from the field data. The estimated effective capacity for either Site-2 or Site-3 lies within about 5% difference with the actual capacity during the incident period.

8.3.4. Model reliability

Note that the parameter values for the merging ratio and the variance of acceleration rates are set with information from the literature [136, 138], because such data are extremely difficult to collect at the acceptable quality under affordable costs. Hence, extensive sensitivity analyses of the proposed model's outputs with respect to those two parameters have been conducted, and shown in Figure 45.

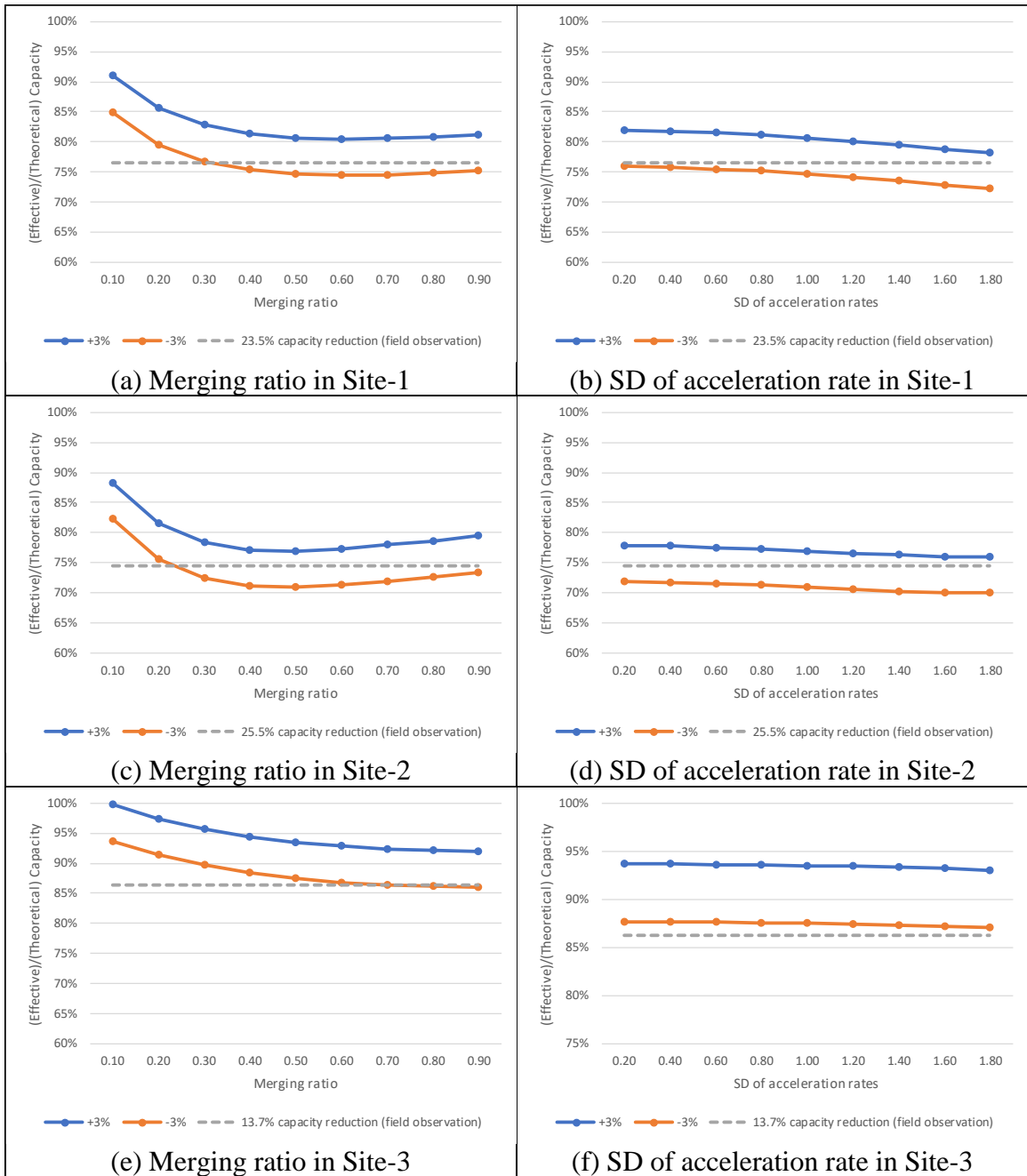


Figure 45. Effective capacity with different merging ratio and SD of acceleration rate in each incident scene

As shown in Figures 45 (a) and (c), the actually observed capacities in the incident scene are 23.5% and 25.5%, respectively, less than the theoretical capacity in Site-1 and Site-2. And, the effective capacities estimated with the developed model in Site-1 and Site-

2 both vary within only $\pm 3\%$ when their merging ratios exceed 0.3. The estimated effective capacity of Site-3 is slightly underestimated, but remains within the comparable range with other sites. Hence, one can conclude that the effective capacity estimated with the developed model is reasonably robust, especially when the merging ratio exceeds 0.3.

The results from Figures 45 (b), (d) and (f), further confirm that the effective capacity estimated with the developed model varies only within 5% (i.e., $\pm 3\%$ at Sites 1 and 2) as long as the standard deviations of acceleration rates lie between 0.2 and 2.0. Such numerical results also support the robust properties of the proposed model for capacity reduction, especially regarding the variance of accelerate rates in the traffic flows.

8.4. Interim findings

In this chapter, the single-lane-open model for estimating the additional capacity reduction due to an incident has been extended to the multi-lane-open model, considering heterogeneous traffic flow properties. In such multi-lane-open scenarios, two types of lane-changing behaviors (i.e., mandatory and discretionary lane changes) are described in the middle-open and rightmost/leftmost-open lanes, and different types of vehicles (i.e., passenger cars and heavy vehicles) along with their heterogeneous properties (i.e., merging times and acceleration rates) are accounted in the model formulations. Lastly, the model developed in this chapter has been analyzed to investigate its properties with respect to robustness and reliability using the reference scenario in the literature and the information collected from the RITIS.

The research findings from this chapter are summarized below:

- The developed multi-lane-open model for estimating the capacity drop due to an incident. one can serve as an effective tool to track the mechanism and relations between the capacity drop and its contributing factors.
- The effective capacity of the middle-open lane is less than that of the rightmost or leftmost lane due to the double-lane occupied process.
- The relatively low acceleration rate of the heavy vehicles and additional interactions resulting from the heterogeneity of traffic flows result in around 3~4% additional capacity reduction, because those factors often contribute to increasing the size of created voids in the traffic flows.

- The effective capacity decreases as the merging ratio increases, but it maintains at the stable level when the merging ratio is over 0.5.
- The effective capacity significantly decreases as the percentage of heavy vehicles increases due to the constraints of their acceleration capability.
- The effective capacity slightly decreases as the standard deviation of the acceleration rate increases, but not at the significant level.

In additions, from the experimental assessment of the developed model using RITIS, with respect to its reliability and robustness, one can have the following findings:

- The differences between the estimates from the developed model and the observed capacity-reduction from the field data are less than 5%.
- The changes in the estimated effective capacities with different merging ratios are less than $\pm 3\%$ for merging ratios exceeding 0.3 (i.e., not very early merging conditions).
- The effective capacity from the developed model varies only within 5% as long as the standard deviations of acceleration rates are between 0.2 and 2.0.

In conclusion, the developed capacity-drop model can yield a reliable and robust estimate, even though precise information about the merging ratio and the standard deviation of acceleration rates may not be available for field applications.

Chapter 9. Assessment of incident impacts and conclusions

9.1. Introduction

To content with traffic incidents and their resulting impacts, highway agencies often need to deploy various cost-efficient traffic control strategies in real time at different phases of incident response and management, based on the estimated incident clearance duration and traffic queue length. However, other than the incident duration, there are various factors that may affect the resulting queues and traffic delays, such as lane-closure status and configuration, capacity drop, traffic volume, incident location, on/off ramp volume, percentage of heavy vehicles, incident severity, rubbernecking effect, weather conditions, geometry, etc. Among those contributing factors, the following three often have the most pronounced impacts on the resulting queues and delays: 1) incident duration, 2) effective capacity (remaining capacity from the lane closure), and 3) incoming traffic volume.

Note that the capacity-drop model developed in Chapters 7 and 8 offers an effective and convenient tool for highway agencies to estimate the available roadway capacity under various conditions. For example, traffic volume from an on/off ramp near an incident scene can directly affect the capacity drop at the lane-closure spot and consequently involve the resulting traffic queues. Such impacts can be captured with the proposed capacity-drop model by changing its merging ratios on the roadway segment. Also, if on-and off-ramps are far from the incident scene, their impacts on the capacity drop, estimable from the proposed model, may be insignificant based on its computed effective merging space.

There are various methods available in the literature for estimate of the queue length and delay, such as statistical approaches, simulation-based models, and analytical approximation. As described in the previous chapters, the statistical and simulation-based

approaches cannot capture the complex relationships among those key factors contributing to incident impacts in a trackable format for the responsible agencies to identify cost-effective control strategies in the time-varying traffic environment. For instance, lane-closure status and incoming traffic volume from the upstream segment of an incident scene may vary over time during different phases of incident management due to sequential implementation of various control strategies. In practice, a response team usually needs to tow vehicles involved in accidents to a shoulder lane so as to minimize the resulting traffic conflicts due to the lane closure. Depending on the severity level, the incident response team may choose to secure more space for safety or efficient clearance process. Hence, such statistical and simulation-based models, based on off-line calibrated data, are not sufficiently flexible in dealing with those time-varying traffic conditions in real-time operations.

This chapter will present a process, based on key models developed in previous chapters, for estimating the traffic queues and delays in real-time clearance and management of a detected highway incident.

9.2. Assessment of incident impacts

Two different analytical methods (i.e., queuing theory and shockwave analysis) are commonly used to estimate queue propagation and its resulting delay. The queuing analysis methods are often employed to compute traffic queues, based on the difference between the total arrival and departure traffic flow rates [152]. The shockwave methodology offers an analytical way to capture traffic queues, based on the changes in traffic density and flow rate along a roadway segment [152]. The main difference between two methods lies in that

the former can yield the estimate of only vertical queues without considering their spatial dimension, but the latter takes into account the temporal and spatial evolution of traffic queues [152, 153]. Park [154] in his study has demonstrated that the shockwave analysis can better describe traffic queue propagation in the scenario where congestion is caused by varying departure capacities at the lane-closure highway segment.

9.2.1. Queue distance

A maximum queue distance is defined as the farthest distance to which the shockwaves due to the lane-closure can propagate over the time periods of incident clearance and traffic recovery. Figure 46 shows the queue evolution diagram over different phases $\{j = 0 \dots n\}$ for incident response, clearance, and traffic recovery. Table 32 shows the notations of variables used for illustrating the time-varying traffic queue estimation process.

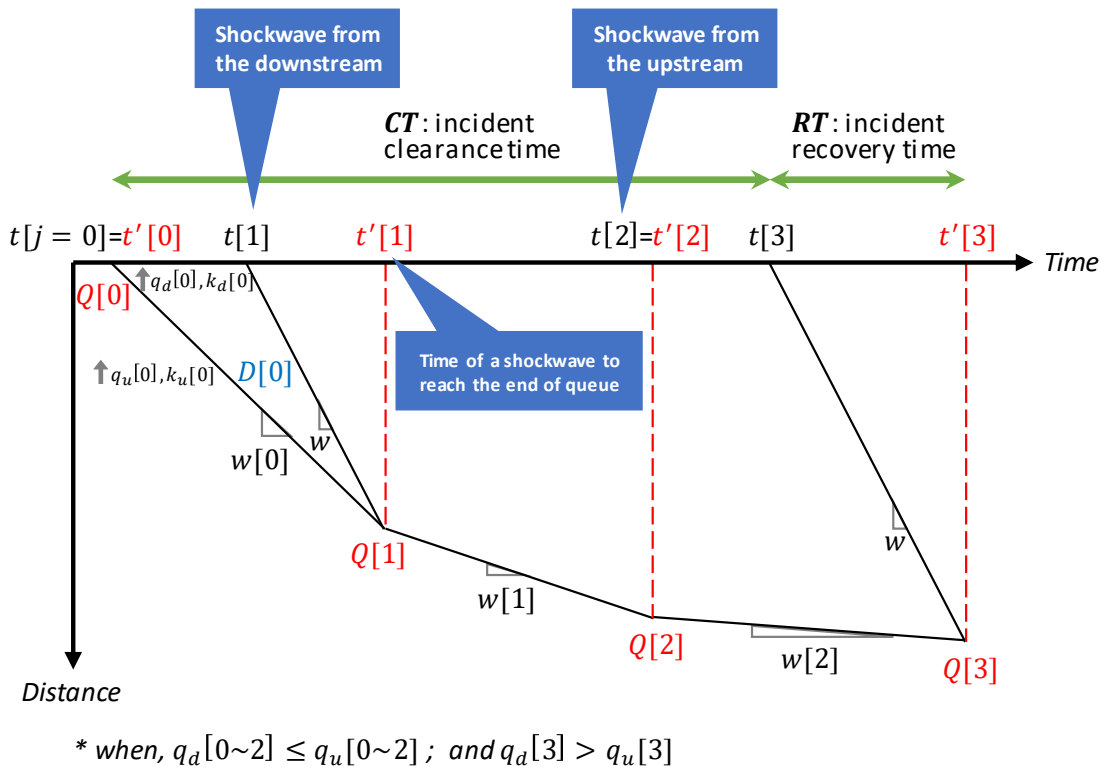


Figure 46. Queuing diagram through three phases

Table 32. Notations for time-varying traffic queue estimation process

Variable	Definition
j	j : a phase of incidents, and 'j = 0' when incident occurred. The last phase (i.e., $j = n$) is defined as when a shockwave by phase j reaches the end of the queue and such queue is completely disappeared.
$t[j]$	$t[j]$: starting time of phase j
$t'[j]$	$t'[j]$: time when a shockwave during phase j reaches the end of queue
$q_d[j]$	$q_d[j]$: outgoing flow from the downstream of a queue (i.e., effective capacity, $C^e[j]$, in the downstream of the lane-closure spot)
$q_u[j]$	$q_u[j]$: incoming flow from the upstream of a queue
$k_d[j]$	$k_d[j]$: density of the outgoing flow
$k_u[j]$	$k_u[j]$: density of incoming flow
$w[j]$	$w[j]$: shockwave speed by a phase j
w	w : shockwave speed between the congested conditions in the triangular fundamental diagram
$Q[j]$	$Q[j]$: queue distance of a phase j , and 'Q[0] = 0'
$D[j]$	$D[j]$: total delay of a phase j

For example, several travel lanes of a roadway may be blocked by an incident at the time, $t[0]$, and the queue will propagate from the blockage scene to its upstream segments at a shockwave speed (i.e., $w[0]$). After the response team arrives at the incident site, some of the blocked travel lanes may be first reopened (i.e., $t[1]$), and thus create a new shockwave from the incident site to catch the first queue wave, and finally form a new queue wave at a new propagation speed (i.e., $w[1]$). Note that the incoming flow rate that affects the speed of queue propagation (i.e., $q_u[j]$) may also vary with the deployed control strategy, such as detour operations and drivers' changes of routes due to the expected excessive delay. When the incident is cleared at the time $t[3]$, the queues will start to dissipate from the lane closure scene and then completely vanish at the time $t'[3]$. Thus, the incident clearance time is from $t[0]$ to $t[3]$, and the recovery time is from $t[3]$ to $t'[3]$.

The shockwave speed for queue formation can be computed from the difference between two adjacent traffic conditions as shown in Equation (23). Note that the outgoing

traffic flow rate in the downstream segment, $q_d[j]$, is approximately equal to the effective capacity, $C^e[j]$, at the lane-closure spot. Such a shockwave speed under the congested traffic condition can be computed with Equation (23) using the theoretical capacity (C^t), free-flow speed (u_f), and jam density (k_j).

$$w[j] = \frac{q_d[j] - q_u[j]}{k_d[j] - k_u[j]}, \quad w = \frac{C^t}{\frac{C^t}{u_f} - k_j} \quad (23)$$

The time for a shockwave at phase j to reach the end of the queues can be computed with a queue length, $Q[j]$, and the shockwave speed (w) as shown in Equation (24). Note that the shockwave from the upstream, due to the incoming flow changes (i.e., $q_u[j] \neq q_u[j - 1]$), immediately affects propagation of the queues in the upstream segment. Thus, $t'[j]$ is the same as $t[j]$.

$$t'[j] = \begin{cases} t[j] + \left(\frac{Q[j]}{w}\right) & \text{if } q_d[j] \neq q_d[j - 1] \\ t[j] & \text{if } q_u[j] \neq q_u[j - 1] \end{cases} \quad (24)$$

Therefore, the queue length in each phase can be estimated with Equation (25). The logic behind the Equation (25) is that a queue distance at each phase (i.e., $Q[j]$) can be computed by adding the previous queue distance (i.e., $Q[j - 1]$) to the increased queue length with its shockwave speed before a new shockwave from phase j reaches the end of the queues (i.e., $(-w[j - 1]) \cdot (t'[j] - t'[j - 1])$).

$$Q[j] = Q[j - 1] + (-w[j - 1]) \cdot (t'[j] - t'[j - 1]) \quad (25)$$

Lastly, the maximum queue distance during the incident clearance period can be computed with Equation (26) as follows:

$$Q_m = \max(Q[j]) \quad (26)$$

9.2.2. Total delay

The total delay is defined as the sum of excessive travel times of vehicles due to a traffic incident, which is a critical measurement for evaluating the performance of incident response and traffic management. Such total delay can be computed with the total queued space-time area at each phase, using the base rule of ‘average height times’ for the triangle area or parallelogram [155], as shown in Equation (27):

$$D = \sum_{j=0}^{n-1} \underbrace{\left\{ \frac{1}{2} (Q[j+1] + Q[j]) \cdot \left(\left(t'[j+1] - \frac{Q[j+1]}{w} \right) - \left(t'[j] - \frac{Q[j]}{w} \right) \right) \right\}}_{\textcircled{1}} \times \underbrace{k_d[j]}_{\textcircled{2}} \times \underbrace{\left(1 - \frac{(q_d[j])}{u_f} \right)}_{\textcircled{3}} \quad (27)$$

The first term, $\textcircled{1}$, of Equation (27) is the total vehicle space-time area of queued vehicles computed with the base rule of the average height times. Then, the total delay can be computed by multiplying this space-time area to the vehicle density during each queue phase (i.e., $\textcircled{2}$), and adjusted with the speed difference between the queuing and free flow conditions (i.e., $\textcircled{3}$).

9.3. Example of system applications

As mentioned in Chapter 1, most highway agencies over the past decades have implemented various incident response plans and comprehensive management systems to contend with non-recurring congestion in highway networks. Such traffic incident management systems are implemented to minimize incident impacts on the transportation network and to recover the traffic conditions as safely and quickly as possible. Ideally, such systems shall be capable of predicting incident duration in real time, estimating limited roadway capacity due to lane closures, and predicting the resulting incident impacts in real time. Using such information, the responsible highway agencies can provide reliable traveler information to drivers on the route, and deploy various cost-effective operational control strategies, including detour plan, early merging, ramp metering, variable speed limit control, etc. However, providing such essential information to the operation center and roadway users in real time with acceptable reliability is quite challenging due to various data and operational issues.

Hence, providing the incident impact information and to support the TIM system's efficient real-time operations, this section presents a real-time incident impact estimation system which consists of the following models:

- Interval-based model for incident duration;
- Point-based model for incident duration;
- Supplemental rules for outliers and integration between the interval and point estimates;
- Capacity-drop estimation model due to lane closures; and
- Time-varying queue-distance estimation process.

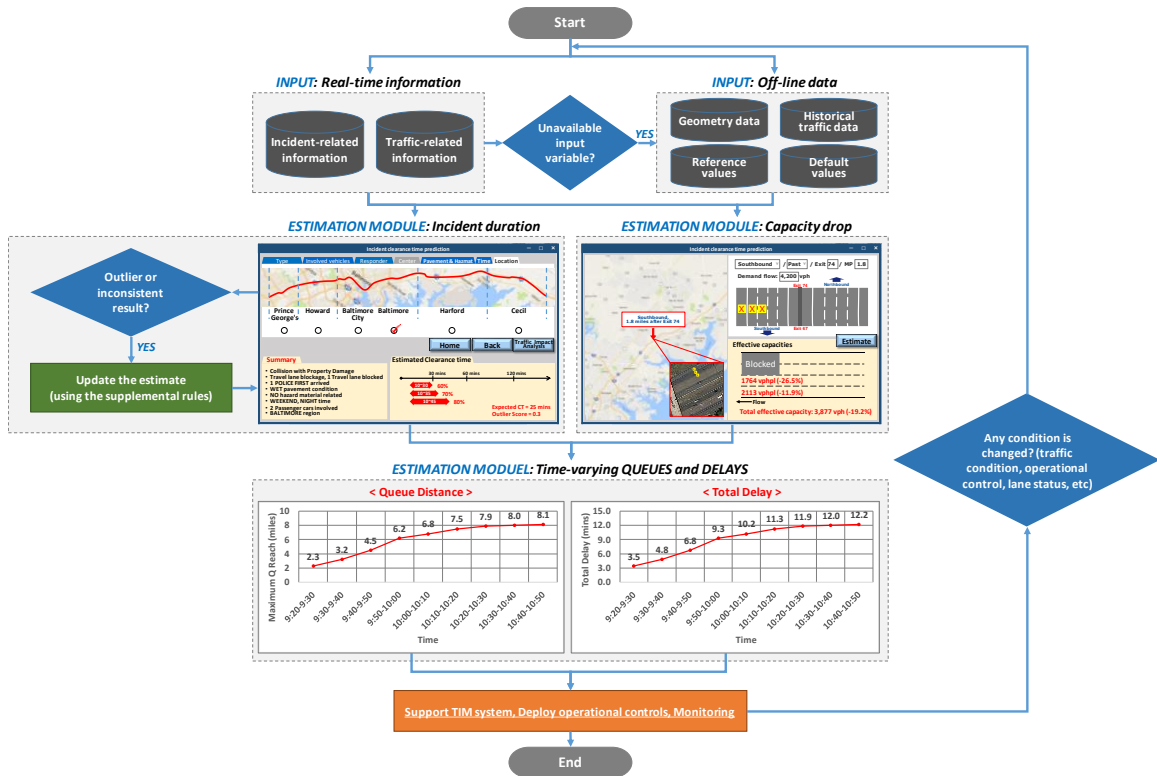


Figure 47. Example of TIM support system application

The process for executing the proposed system is illustrated below:

<Input module: on/off-line information>

- Detect an incident using patrol vehicles (e.g., CHART units, POLICE, etc.), surveillance cameras, and reports from citizen or drivers, etc.
- Start to collect incident-related and traffic-related information from the response team and available traffic devices near the incident scene (e.g., detectors, CCTV, etc.)
- Apply the off-line data system to collect the following information for estimating incident impacts:

- Geometry data: lane configuration, alignment of road, and location of nearest on/off ramp;
- Historical traffic data: traffic volume, density, and speed information (for estimating speed-flow-density relationships);
- Reference/default values: acceleration rate of vehicles, merging time, etc.
- Use a reference/default value of the off-line data system, if any required information is unavailable.

<Estimation Module: incident duration>

- Estimate the duration of a detected incident, based on the information available from the scene and the off-line database including: time, location, incident type, lane-closure status, involved vehicles, response units, operation center, pavement condition, and hazard materials.
- Provide the initial estimate of the incident duration with real-time first available data (i.e., interval estimate with different confidence levels) and then update the estimate when additional information that can better the prediction becomes available sequentially.
- When all required information is available, provide a more precise estimate of the incident duration with the more advanced method such as the ensemble model in Chapter 5.
- If an outlier of incidents is detected (i.e., when an outlier score is greater than the outlier threshold), updated the estimated result using the developed rules from the outlier analysis in Chapter 6.

- Such outlier information can be utilized to assess the system’s performance at the off-line level, including the data recording and inventory process (e.g., data recording and processing errors, missing critical information from the scene, lack of operating resource to deal with a severe (but infrequent) incident, etc.)

<Estimation Module: capacity drop>

- Estimate the available roadway capacity due to a detected incident, based on the incident-and traffic-related information, etc.
 - Lane-closure status, geometry information, historical traffic information (i.e., traffic volume, density, and speed), default values (e.g., acceleration rate of vehicles, merging time, etc.)
- Provide the estimated effective capacity at the lane-closure spot and the effective merging space, a critical area within which vehicles making lane changes may affect the capacity drop on the lane-closure roadway segment.

<Output module: time-varying queue distance and delay>

- If the incoming traffic flow is greater than the available capacity at the lane-closure spot, then proceed the estimate of queue distance over time during each phase of the incident management process under its time-varying traffic conditions.

Required information	Outputs
<ul style="list-style-type: none"> • Incident duration (clearance time) • Available capacity (effective capacity) • Real-time incoming traffic volumes (due to drivers’ route changes and natural traffic flow fluctuation) • Time-varying incident management phases • Deployed control strategies 	<ul style="list-style-type: none"> • Time-varying queue distance • Time-varying total delay

<Supporting TIM center and Monitoring>

- Forward the information of the time-varying incident impacts (queue and travel time delay) to the TIM center to support a decision-making process.
- Compute the total costs with respect to operations, safety, and economy/environment, due to the congestion incurred by a traffic incident.

Operational MOE	Safety MOE	Economy/environment MOE
<ul style="list-style-type: none"> • Travel time / Delay • Queue length 	<ul style="list-style-type: none"> • Speed difference / variances • # of traffic conflicts • Time-to-collision (TTC) 	<ul style="list-style-type: none"> • Fuel consumption • Emission

- Assess all available control strategies using cost-benefit analysis
- Select and implement the most cost-effective control strategies.
- Keep monitoring and updating the traffic conditions on all candidate detouring routes.
- Go back to the first stage of the TIM support system, if any change that may affect the resulting traffic impacts (e.g., traffic conditions, lane status, incident management phase, etc.) has been detected.

9.3.1. Illustration of system applications

This study chooses one incident case from the CHART database, and shows how the proposed TIM support system can be used to estimate the traffic incident impacts. Table 33 shows incident-related information of the selected case from the CHART database. The procedures, shown in Chapter 8, are to first utilize the detector data near the incident scene from the RITIS, and then to estimate the fundamental diagram of traffic flows to calibrate the q-k relations as shown Table 33.

Table 33. Basic input information for the application example

Incident-related information (Sep. 2, 2017)	Incident duration	Location	Lane status
	20:55~21:54 Actual CT: 59 mins	I-95 North, before Exit 49	2 of 4 lanes blocked (right)
	Detail		
	CPI, 3 passenger cars (Overturned), Weekend, Nighttime, Wet pavement condition, 2 CHARTs, 1 POLICE, 1 FIRE		
Traffic-related information (Jul.~Nov., 2016)	Observed maximum flow rate (vphpl)	Free-flow speed (km/h)	Jam density (vpkpl)
	1769	98.0	44.9

This example has also employed the traffic volume data from the upstream detectors of the incident scene to reflect the time-varying incoming traffic flow rate.

Table 34. Incoming traffic flow rate from the detectors near the scene (Sep. 2, 2017)

Time	Incoming traffic flow rate (vph)	Time	Incoming traffic flow rate (vph)
20:45	3605	21:30	2790
21:00	3160	21:45	2890
21:15	2990	22:00	2545

Based on all data and information associated with the selected incident case, one can apply the proposed process to produce the event timetable, as shown in Table 35. Note that

the incident report contains the information of lane status during the operations, such as which and when a lane is closed or reopened.

Table 35. An example of output produced from the proposed incident management system

Event	Time	Detail
①	20:55	<ul style="list-style-type: none"> • Incident detected, 2 of 4 lanes blocked • Initial estimate of incident duration <ul style="list-style-type: none"> ○ 25~60(60%), 10~60(70%), and 10~70(80%) minutes ○ Predicted CT: 50 minutes • Point estimate of incident duration <ul style="list-style-type: none"> ○ Predicted CT: 56 minutes, outlier score: 0.36 • Effective capacity (2 of 4 opened): 2714 vph (76.7%)
②	21:00	• Incoming traffic flow rate changed: 3605 → 3160 vhp
③	21:15	• Incoming traffic flow rate changed: 3160 → 2990 vhp
④	21:17	<ul style="list-style-type: none"> • One more lane blocked (3 of 4 lanes blocked) • Effective capacity (1 of 4 opened): 1435 vph (81.1%)
⑤	21:30	• Incoming traffic flow rate changed: 2990 → 2790 vhp
⑥	21:45	• Incoming traffic flow rate changed: 2790 → 2890 vhp
⑦	21:46	<ul style="list-style-type: none"> • One lane reopened (2 of 4 lanes blocked) • Effective capacity (2 of 4 opened): 2714 vph (76.7%)
⑧	21:54	• Incident cleared; all lanes reopened
⑨	22:00	• Incoming traffic flow rate changed: 2890 → 2545 vhp

As shown in Event-① in Table 35, when an incident blocking two out of the four travel lanes is detected, the system can execute the incident duration module to first provide the initial estimate of incident duration using only available information, such as ‘nighttime’ and ‘no truck or no tow service’ in this example. And then the system can further apply the advanced module of the incident duration module to provide the more precise estimate, when all required information becomes available from the field operations. The estimated incident clearance time of 56 minutes is very close to the actual incident clearance time of 59 minutes in this example.

The effective capacity according to the lane status can be computed by using the traffic-related information from the nearest detectors and the capacity drop estimation model developed in the previous chapters. Some parameters used in this example are from the RITIS and its CCTVs. This is inevitable in the real-world system applications as some essential model variables or parameters often are not available in real-time incident response and clearance process.

9.3.2. Estimation results of the selected case

Figure 48 shows the evolution of queue distance up to each incident phase, based on the formulations presented in this chapter. First, when an incident occurs, the queues start to propagate to its upstream segment, but at a slow pace with its shockwave speed of less than 10km/h. However, when the one additional lane is closed for facilitating field operations at 21:17, the queues have developed rapidly and reached around 8km from the scene before the arrival of another shockwave, generated due to the lane-reopening operations. After those phases, the traffic condition becomes quite stable, and the incident is cleared at 21:54 and fully recovered at 22:02. Lastly, the maximum queue distance is 8.66km when the incoming traffic flow rate is less than the outgoing traffic flow rate, and the total delay is 623.7 hours in this example. Lastly, such time-varying incident impacts can be directly displayed on the map as shown in Figure 49, such information also shows when and which on/off ramp flows in the upstream segment has been interrupted.

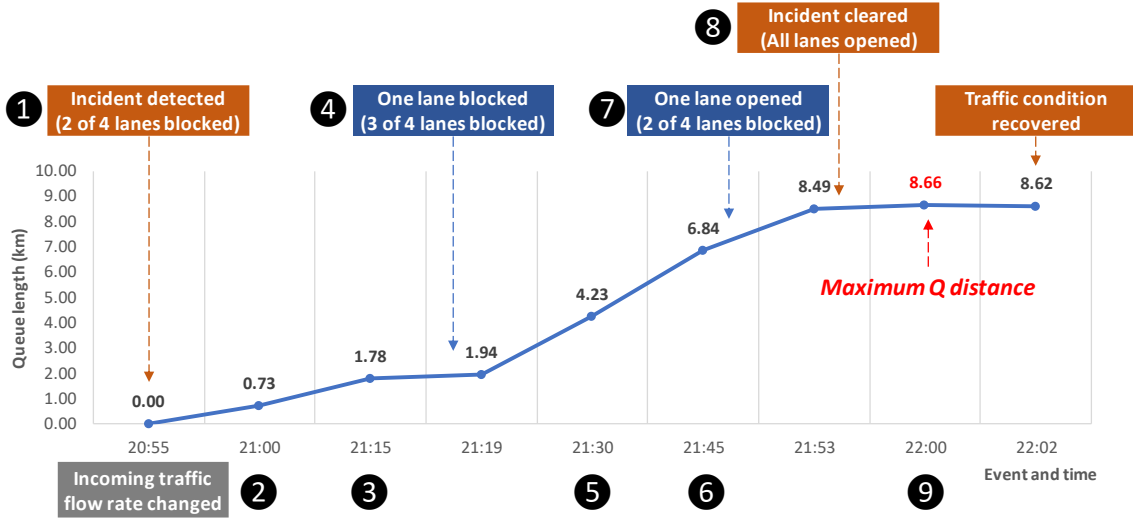


Figure 48. Time-varying queue distance up to each incident phase

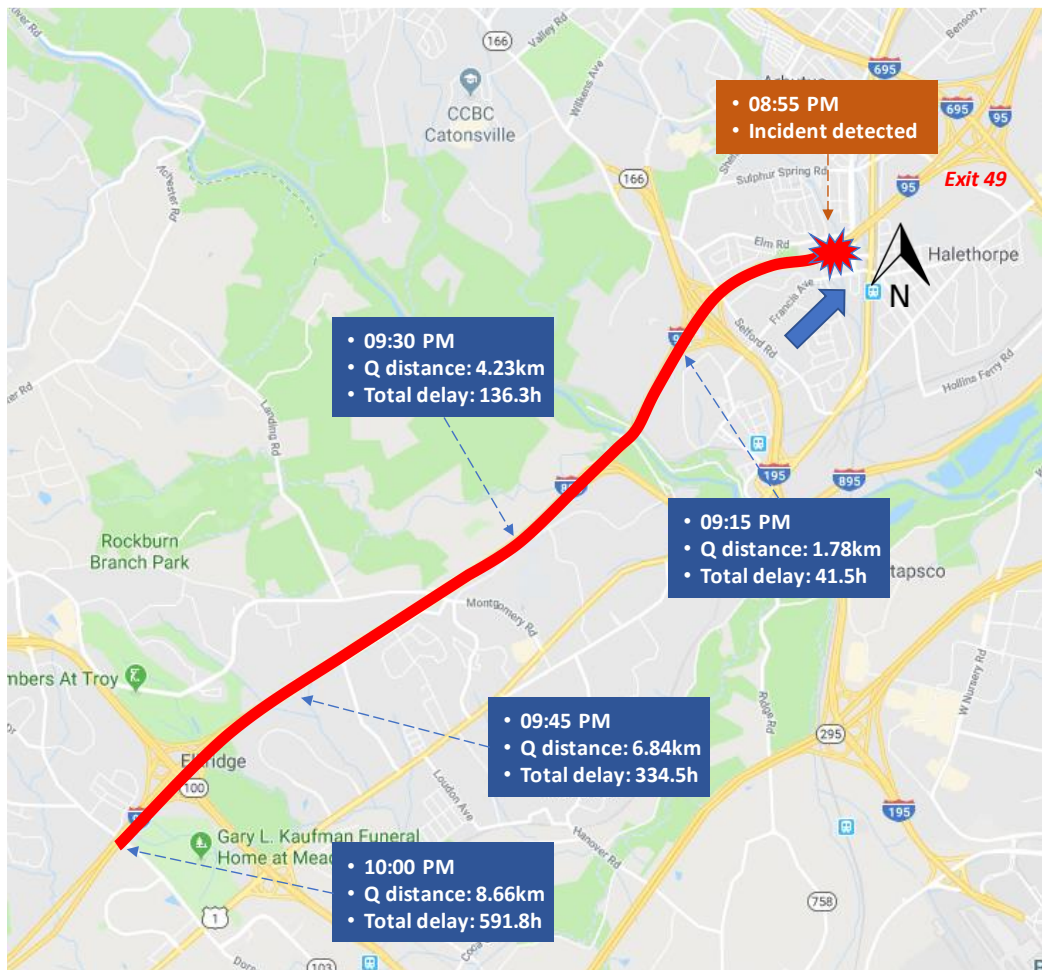


Figure 49. Example of the queue propagation information on the map

9.4. Conclusions and future research

9.4.1. Conclusions

Traffic incidents often cause not only significant delays to daily roadway users, but also degrading reliability of a highway network. To mitigate the impacts of such incidents and to recover the roadway's performance sufficiently efficient to comply with motorists' expectation, it is essential for the responsible highway agencies to have an intelligent support system that can accurately estimate the required clearance duration in real time, assess the resulting traffic impacts such as the time-varying queue length, and then advise necessary control actions.

However, the task for developing a comprehensive and cost-efficient traffic incident management system for real-world operations remains quite challenging due to the following issues:

- **Limited data availability at the early stage of incident responses:** Due to the emergency nature of incident response operations, sufficient information for estimating the incident impacts may not be available in real time or even collected.
- **Unique characteristics and complexities of factors contributing to the incident impacts:** Most incident duration data exhibit a highly skewed distribution and complex correlations among its explanatory variables, which include both qualitative and quantitative factors.
- **No reliable tool is available for estimating critical information:** Due to the unique and complex nature of incident duration data, most developed models or methods are for location-specific applications and cannot be generalized to other locations.

- **Insufficient training for field data collection and database development:** Most highway agencies in response to incidents and traffic management mostly focus on the efficient recovery of roadway conditions, generally paying less attention to collecting all essential data for offline analysis, evaluation, and model development.

To contend with those critical issues, this study has presented a traffic incident management (TIM) support system which includes: 1) an interval-based model for initial estimate of incident duration; 2) a point-based model for a more precise estimate of incident duration; 3) supplemental rules for outlier identification and an integrated estimation process to take advantages of the strengths from both interval and point estimates; 4) a capacity-drop estimation model to compute the available roadway capacity due to lane closures; and 5) time-varying queue and delay estimation models, based on the estimated incident duration and available roadway capacity.

The first component is designed to provide an initial estimate of incident duration at the early stage of incident response and clearance. Such estimates then be updated when more related information becomes available in the real-time incident response process. Also, the proposed TIM system in this study contains an outlier analysis component and integration process for tackling inconsistent estimation results. Using the supplemental rules from those outlier and integration process, the incident duration prediction system can yield more reliable and accurate estimation results.

The second key component of the proposed TIM support system is for estimating the additional capacity reduction (i.e., capacity drop) due to a detected incident and its lane-closure operations. This capacity-drop model also provides an estimate of the effective

merging space, defined as the segment of a highway measured from the lane-closure spot, within which a merging maneuver taking place will cause the capacity drop.

The last component is for assessing the time-varying queues and delays due to the incident response and clearance, depending on all contributing factors such as the sequence of the incident management tasks, implementation of operational strategies, drivers' route changes, and fluctuation of traffic volume.

9.4.2. Contributions

The TIM support system proposed in this study has made the following contributions for use in real-time incident response and management:

- Providing an initial estimate with different confidence levels for a detected incident, based on the limited information at the early stage of incident responses.
- Presenting a specially-design estimation process to circumvent the complexity embedded in the nature of incident data, such as highly skewed distribution, complex correlations among the explanatory variables, different significant-levels according to the different incident conditions, mixed qualitative and quantitative variables, and heteroscedasticity.
- Developing an innovative estimating process to yield the robust point estimates, which can be used to improve the performance of any different primary model used for predicting incident duration.
- Estimating the level of outlier-ness associated with each incident, and offering a tool for detecting anomalies and noises.

- Improving the incident duration estimate using the supplemental rules developed from the outlier analysis.
- Proposing an integrated estimation process to take advantage of the strengths from the point-based and interval-based models.
- Developing an analytical capacity-drop estimation model to track the critical relationships between the key contributing factors and the resulting available capacity that can facilitate the control center in better making real-time decisions.
- Producing a set of formulations for computing the effective merging space that offers the basis for field incident respondents to design a proper merge control.

9.4.3. Future research

Although this study has presented a convenient and cost-efficient TIM support system, the following issues remain to be done to ensure its effective operations in practice:

- **Model transferability:** Because the incident duration prediction model in this study was developed by using the data of I-95 in Maryland, its transferability to other highways remains to be evaluated, and a rigorous process and model for updating key parameters, given more new data, shall also be developed.
- **Utilization of the outlier information for system performance assessment:** In the outlier analysis process, one can use the proposed procedures to detect noises or faulty data associated with incidents. Such information can be further utilized to evaluate the quality of incident-data management team and assess if the resources

and training are adequate or to support the efficient incident response and traffic management operations.

- **Location of the incident and on/off ramp:** The effective capacity can vary with the location of a detected incident and its distance to the nearest on/off ramp, because such a relationship may change the merging ratio which is one of the key factors for the capacity drop. For both convenience and reliability concerns, one shall extend the proposed capacity-drop model to incorporate explicitly its relationship with the near on-and-off-ramp distance and flow rates.
- **Prediction of time-varying incoming traffic volume:** As mentioned in this chapter, the incoming traffic volume is one of the critical input values to estimate time-varying incident impacts. The variation in the incoming traffic volume can be caused by various factors such as drivers' route choice behaviors, implementation of traffic control strategies, nature fluctuation of flows, etc. Therefore, development of a reliable model for producing such information in real time should be one essential task.

Bibliography

1. Schrank., D., et al., *2015 Urban Mobility Scorecard*. 2015.
2. FHWA. *Operations Story*. 2017 [cited 2018 August 15]; Available from: <https://ops.fhwa.dot.gov/aboutus/opstory.htm>.
3. Kim, W. *Development of a traffic incident management system for contending with non-recurrent highway congestion*. in *Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR)*. 2014. IEEE Computer Society.
4. Kim, W., et al., *Development of a Traffic Management Decision Support Tool for Freeway Incident Traffic Management (FITM) Plan Deployment*. 2017, Maryland. State Highway Administration.
5. Hojati, A.T., et al., *Modelling total duration of traffic incidents including incident detection and recovery time*. *Accident Analysis & Prevention*, 2014. **71**: p. 296-305.
6. Al Kaabi, A., D. Dissanayake, and R. Bird, *Response time of highway traffic accidents in Abu Dhabi: Investigation with hazard-based duration models*. *Transportation research record*, 2012. **2278**(1): p. 95-103.
7. Nam, D. and F. Mannering, *An exploratory hazard-based analysis of highway incident duration*. *Transportation Research Part a-Policy and Practice*, 2000. **34**(2): p. 85-102.
8. Li, R., *Traffic incident duration analysis and prediction models based on the survival analysis approach*. *IET intelligent transport systems*, 2014. **9**(4): p. 351-358.
9. Li, R., F.C. Pereira, and M.E. Ben-Akiva, *Overview of traffic incident duration analysis and prediction*. *European Transport Research Review*, 2018. **10**(2): p. 22.
10. Won, M., H. Kim, and G.-L. Chang, *Knowledge-Based System for Estimating Incident Clearance Duration for Maryland I-95*. *Transportation Research Record: Journal of the Transportation Research Board*, 2018.
11. Golob, T.F., W.W. Recker, and J.D. Leonard, *An analysis of the severity and incident duration of truck-involved freeway accidents*. *Accid Anal Prev*, 1987. **19**(5): p. 375-95.
12. Giuliano, G., *Incident characteristics, frequency, and duration on a high volume urban freeway*. *Transp Res A*, 1989. **23**.
13. Chung, Y. and B.J. Yoon, *Analytical method to estimate accident duration using archived speed profile and its statistical analysis*. *KSCE J Civ Eng*, 2012. **16**.
14. Zhang, H. and A.J. Khattak, *Analysis of cascading incident event durations on urban freeways*. *Transp Res Rec*, 2010. **2178**.
15. Hojati, A.T., et al., *Analysing freeway traffic incident duration using an Australian data set*. *Road Transp Res*, 2012. **21**.
16. Chimba, D., et al., *Impact of abandoned and disabled vehicles on freeway incident duration*. *Journal of transportation engineering*, 2013. **140**(3): p. 04013013.
17. Jones, B., L. Janssen, and F. Mannering, *Analysis of the frequency and duration of freeway accidents in Seattle*. *Accid Anal Prev*, 1991. **23**(4): p. 239-55.
18. Junhua, W., C. Haozhe, and Q. Shi, *Estimating freeway incident duration using accelerated failure time modeling*. *Safety Science*, 2013. **54**: p. 43-50.

19. Alkaabi, A.M.S., D. Dissanayake, and R. Bird, *Analyzing clearance time of urban traffic accidents in Abu Dhabi, United Arab Emirates, with hazard-based duration modeling method*. Transp Res Rec, 2011. **2229**.
20. Ghosh, I., P.T. Savolainen, and T.J. Gates, *Examination of factors affecting freeway incident clearance times: a comparison of the generalized F model and several alternative nested models*. J Adv Transport, 2014. **48**.
21. Hojati, A.T., et al., *Hazard based models for freeway traffic incident duration*. Accid Anal Prev, 2013. **52**.
22. Zou, Y., et al., *Application of finite mixture models for analysing freeway incident clearance time*. Transportmetrica A: Transport Science, 2016. **12**(2): p. 99-115.
23. Khattak, A.J., H.M. Al-Deek, and R.W. Hall, *Concept of an advanced traveler information system testbed for the Bay area: research issues*. Journal of Intelligent Transportation Systems, 1994. **2**(1): p. 45-71.
24. Khattak, A.J., J.L. Schofer, and M.-h. Wang, *A simple time sequential procedure for predicting freeway incident duration*. Journal of Intelligent Transportation Systems, 1995. **2**(2): p. 113-138.
25. Garib, A., A.E. Radwan, and H. Al-Deek, *Estimating magnitude and duration of incident delays*. Journal of Transportation Engineering, 1997. **123**(6): p. 459-466.
26. El-Basyouny, K. and T. Sayed, *Comparison of two negative binomial regression techniques in developing accident prediction models*. Transportation Research Record: Journal of the Transportation Research Board, 2006(1950): p. 9-16.
27. Qi, Y. and H. Teng, *An information-based time sequential approach to online incident duration prediction*. Journal of Intelligent Transportation Systems, 2008. **12**(1): p. 1-12.
28. Chung, Y., *Development of an accident duration prediction model on the Korean Freeway Systems*. Accident Analysis & Prevention, 2010. **42**(1): p. 282-289.
29. Khattak, A., X. Wang, and H. Zhang, *Incident management integration tool: dynamically predicting incident durations, secondary incident occurrence and incident delays*. IET Intelligent Transport Systems, 2012. **6**(2): p. 204-214.
30. Wang, X., S. Chen, and W. Zheng, *Traffic incident duration prediction based on partial least squares regression*. Procedia-Social and Behavioral Sciences, 2013. **96**: p. 425-432.
31. Li, R. and P. Shang, *Incident duration modeling using flexible parametric hazard-based models*. Computational intelligence and neuroscience, 2014. **2014**: p. 33.
32. Khattak, A.J., et al., *Modeling Traffic Incident Duration Using Quantile Regression*. Transportation Research Record: Journal of the Transportation Research Board, 2016(2554): p. 139-148.
33. Li, X., A.J. Khattak, and B. Wali, *Large-Scale Traffic Incident Duration Analysis: The Role of Multi-agency Response and On-Scene Times*. 2017.
34. Zou, Y., et al., *Quantile analysis of factors influencing the time taken to clear road traffic incidents*. Proceedings of the Institution of Civil Engineers - Transport, 2017.
35. Peeta, S., J.L. Ramos, and S. Gedela, *Providing Real-Time Traffic Advisory and Route Guidance to Manage Borman Incidents On-Line Using the Hoosier Helper Program*. Joint Transportation Research Program, 1284 Civil Engineering Building, Purdue University, West Lafayette, Indiana 47907-1284. 2000.

36. Yu, B. and Z. Xia, *A methodology for freeway incident duration prediction using computerized historical database*, in *CICTP 2012: Multimodal Transportation Systems - Convenient, Safe, Cost-Effective, Efficient - Proceedings of the 12th COTA International Conference of Transportation Professionals*. 2012.
37. Royston, P. and M.K.B. Parmar, *Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects*. *Statistics in medicine*, 2002. **21**(15): p. 2175-2197.
38. Royston, P. and P.C. Lambert, *Flexible parametric survival analysis using Stata: beyond the Cox model*. 2011.
39. Ozbay, K. and P. Kachroo, *Incident management in intelligent transportation systems*. 1999.
40. Ozbay, K. and N. Noyan, *Estimation of incident clearance times using Bayesian Networks approach*. *Accident Analysis & Prevention*, 2006. **38**(3): p. 542-555.
41. Boyles, S., D. Fajardo, and S.T. Waller. *A naive Bayesian classifier for incident duration prediction*. in *86th Annual Meeting of the Transportation Research Board, Washington, DC*. 2007.
42. Xiaoqiang, Z., L. Ruimin, and Y. Xinxin. *Incident duration model on urban freeways based on Classification and Regression Tree*. in *Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on*. 2009. IEEE.
43. Chang, H.-L. and T.-P. Chang, *Prediction of freeway incident duration based on classification tree analysis*. *Journal of the Eastern Asia Society for Transportation Studies*, 2013. **10**: p. 1964-1977.
44. Ma, X., et al., *Prioritizing Influential Factors for Freeway Incident Clearance Time Prediction Using the Gradient Boosting Decision Trees Method*. *IEEE Transactions on Intelligent Transportation Systems*, 2017.
45. Zhan, C., A. Gan, and M. Hadi, *Prediction of lane clearance time of freeway incidents using the M5P tree algorithm*. *IEEE Transactions on Intelligent Transportation Systems*, 2011.
46. Smith, K. and B.L. Smith, *Forecasting the Clearance Time of Freeway Accidents*. 2001, Charlottesville: Center for Transportation Studies, University of Virginia.
47. Wang, W., H. Chen, and M.C. Bell, *Vehicle breakdown duration modelling*. *Journal of Transportation and Statistics*, 2005. **8**(1): p. 75-84.
48. Wei, C.-H. and Y. Lee, *Sequential forecast of incident duration using Artificial Neural Network models*. *Accident Analysis & Prevention*, 2007. **39**(5): p. 944-954.
49. Valenti, G., M. Lelli, and D. Cucina, *A comparative study of models for the incident duration prediction*. *European Transport Research Review*, 2010. **2**(2): p. 103-111.
50. Guan, L., et al. *Traffic incident duration prediction based on Artificial Neural Network*. in *Intelligent Computation Technology and Automation (ICICTA), 2010 International Conference on*. 2010. IEEE.
51. Wu, W.-w., S.-y. Chen, and C.-j. Zheng, *Traffic incident duration prediction based on support vector regression*, in *ICCTP 2011: Towards Sustainable Transportation Systems*. 2011. p. 2412-2421.

52. Vlahogianni, E.I. and M.G. Karlaftis, *Fuzzy-Entropy Neural Network Freeway Incident Duration Modeling with Single and Competing Uncertainties*. Computer-Aided Civil and Infrastructure Engineering, 2013. **28**(6): p. 420-433.
53. Park, H., A. Haghani, and X. Zhang, *Interpretation of Bayesian neural networks for predicting the duration of detected incidents*. Journal of Intelligent Transportation Systems, 2016. **20**(4): p. 385-400.
54. Park, H.J., S. Gao, and A. Haghani, *Sequential Interpretation and Prediction of Secondary Incident Probability in Real Time*. 2017.
55. Lin, P.-W., N. Zou, and G.-L. Chang. *Integration of a discrete choice model and a rule-based system for estimation of incident duration: a case study in Maryland*. in *CD-ROM of Proceedings of the 83rd TRB Annual Meeting, Washington, DC*. 2004.
56. Kim, W., S. Natarajan, and G.-L. Chang. *Empirical analysis and modeling of freeway incident duration*. in *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*. 2008. IEEE.
57. Kim, W. and G.-L. Chang, *Development of a Hybrid Prediction Model for Freeway Incident Duration: A Case Study in Maryland*. International Journal of Intelligent Transportation Systems Research, 2011. **10**(1): p. 22-33.
58. Ji, Y., *Prediction of freeway incident duration based on the Multi-model fusion algorithm*. 2011: p. 3327--3329.
59. He, Q., et al., *Incident duration prediction with hybrid tree-based quantile regression*, in *Advances in Dynamic Network Modeling in Complex Transportation Systems*. 2013, Springer. p. 287-305.
60. Li, R., F.C. Pereira, and M.E. Ben-Akiva, *Competing risk mixture model and text analysis for sequential incident duration prediction*. Transportation Research Part C: Emerging Technologies, 2015. **54**: p. 74-85.
61. Lin, L., Q. Wang, and A.W. Sadek, *A combined M5P tree and hazard-based duration model for predicting urban freeway traffic accident durations*. Accid Anal Prev, 2016. **91**.
62. Zhu, W., et al., *Predicting Incident Duration based on Spatiotemporal Heterogeneous Pattern Recognition*. 2017.
63. Aggarwal, C.C., *Outlier analysis*. 2013.
64. Krammes, R.A. and G.O. Lopez, *Updated capacity values for short-term freeway work zone lane closures*. Transportation Research Record, 1994.
65. Al-Kaisy, A., M. Zhou, and F. Hall, *New Insights into Freeway Capacity at Work Zones: Empirical Case Study*. Transportation Research Record: Journal of the Transportation Research Board, 2000. **1710**: p. 154-160.
66. Kim, T., D.J. Lovell, and J. Paracha. *A new methodology to estimate capacity for freeway work zones*. in *80th Annual Meeting of the Transportation Research Board, Washington, DC*. 2001.
67. Al-Kaisy, A. and F. Hall, *Guidelines for Estimating Capacity at Freeway Reconstruction Zones*. Journal of Transportation Engineering, 2003. **129**(5): p. 572-577.
68. Benekohal, R., A.-Z. Kaja-Mohideen, and M. Chitturi, *Methodology for Estimating Operating Speed and Capacity in Work Zones*. Transportation Research Record: Journal of the Transportation Research Board, 2004. **1883**: p. 103-111.

69. Sarasua, W., et al., *Estimating Interstate Highway Capacity for Short-Term Work Zone Lane Closures: Development of Methodology*. Transportation Research Record: Journal of the Transportation Research Board, 2006. **1948**: p. 45-57.
70. Racha, S., et al., *Analysis of Work Zone Traffic Behavior for Planning Applications*. Transportation Planning and Technology, 2008. **31**(2): p. 183-199.
71. Avrenli, K.A., R. Benekohal, and H. Ramezani, *Determining the Speed-Flow Relationship and Capacity for a Freeway Work Zone*. TRB 2011 Annual Meeting, 2011.
72. Adeli, H. and X. Jiang, *Neuro-Fuzzy Logic Model for Freeway Work Zone Capacity Estimation*. Journal of Transportation Engineering, 2003. **129**(5): p. 484-493.
73. Karim, A. and H. Adeli, *Radial Basis Function Neural Network for Work Zone Capacity and Queue Estimation*. Journal of Transportation Engineering, 2003. **129**(5): p. 494-503.
74. Weng, J. and Q. Meng, *Decision Tree-Based Model for Estimation of Work Zone Capacity*. Transportation Research Record: Journal of the Transportation Research Board, 2011. **2257**: p. 40-50.
75. Zheng, N., et al. *A comparison of freeway work zone capacity prediction models*. in *Procedia - Social and Behavioral Sciences*. 2011.
76. Weng, J. and Q. Meng, *Ensemble Tree Approach to Estimating Work Zone Capacity*. Transportation Research Record: Journal of the Transportation Research Board, 2012. **2286**: p. 56-67.
77. Ping, W.V. and K. Zhu. *Evaluation of work zone capacity estimation models: A computer simulation study*. in *Sixth Asia-Pacific Transportation Development Conference, 19 th ICTPA Annual Meeting, Hong Kong*. 2006.
78. Chatterjee, I., et al., *Replication of Work Zone Capacity Values in a Simulation Model*. Transportation Research Record: Journal of the Transportation Research Board, 2009. **2130**: p. 138-148.
79. Heaslip, K., et al., *Estimation of Freeway Work Zone Capacity Through Simulation and Field Data*. Transportation Research Record: Journal of the Transportation Research Board, 2009. **2130**: p. 16-24.
80. Heaslip, K., M. Jain, and L. Elefteriadou, *Estimation of arterial work zone capacity using simulation*. Transportation Letters, 2011. **3**(2): p. 123-134.
81. Papageorgiou, M., et al., *Real-time merging traffic control with applications to toll plaza and work zone management*. Transportation Research Part C: Emerging Technologies, 2008.
82. Wong, G.C.K. and S.C. Wong, *A multi-class traffic flow model - An extension of LWR model with heterogeneous drivers*. Transportation Research Part A: Policy and Practice, 2002.
83. Treiber, M., A. Kesting, and D. Helbing, *Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps*. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 2006.
84. Nishinari, K., M. Treiber, and D. Helbing, *Interpreting the wide scattering of synchronized traffic data by time gap statistics*. Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, 2003.
85. Zhang, H.M. and T. Kim, *A car-following theory for multiphase vehicular traffic flow*. Transportation Research Part B: Methodological, 2005.

86. Yeo, H., *Asymmetric microscopic driving behavior theory*. 2008: University of California, Berkeley.
87. Oh, S. and H. Yeo, *Impact of stop-and-go waves and lane changes on discharge rate in recovery flow*. Transportation Research Part B: Methodological, 2015.
88. Chen, D., et al., *Microscopic traffic hysteresis in traffic oscillations: A behavioral perspective*. Transportation Research Part B: Methodological, 2012.
89. Chen, D., et al., *On the periodicity of traffic oscillations and capacity drop: The role of driver characteristics*. Transportation Research Part B: Methodological, 2014.
90. Laval, J.A. and C.F. Daganzo, *Lane-changing in traffic streams*. Transportation Research Part B: Methodological, 2006.
91. Duret, A., J. Bouffier, and C. Buisson, *Onset of Congestion from Low-Speed Merging Maneuvers Within Free-Flow Traffic Stream*. Transportation Research Record: Journal of the Transportation Research Board, 2010.
92. Coifman, B. and S. Kim, *Extended bottlenecks, the fundamental relationship, and capacity drop on freeways*. Transportation Research Part A: Policy and Practice, 2011.
93. Leclercq, L., J.A. Laval, and N. Chiabaut, *Capacity drops at merges: An endogenous model*. Transportation Research Part B: Methodological, 2011.
94. Leclercq, L., et al., *Capacity drops at merges: New analytical investigations*. Transportation Research Part C: Emerging Technologies, 2016.
95. Weng, J. and Q. Meng, *Estimating capacity and traffic delay in work zones: An overview*. Transportation Research Part C: Emerging Technologies, 2013.
96. Yuan, K., V.L. Knoop, and S.P. Hoogendoorn, *A Microscopic Investigation Into the Capacity Drop: Impacts of Longitudinal Behavior on the Queue Discharge Rate*. Transportation Science, 2017.
97. Han, Y., D. Chen, and S. Ahn, *Variable speed limit control at fixed freeway bottlenecks using connected vehicles*. Transportation Research Part B: Methodological, 2017. **98**: p. 113-134.
98. Hall, F.L. and K. Agyemang-Duah, *Freeway capacity drop and the definition of capacity*. Transportation research record, 1991(1320).
99. Banks, J.H., *The two-capacity phenomenon: some theoretical issues*. Transportation Research Record, 1991(1320).
100. Kerner, B.S., *Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks*. Physical Review E, 2002. **65**(4): p. 046138.
101. Cassidy, M.J. and R.L. Bertini, *Some traffic features at freeway bottlenecks*. Transportation Research Part B: Methodological, 1999. **33**(1): p. 25-42.
102. Sarvi, M., M. Kuwahara, and A. Ceder, *Observing freeway ramp merging phenomena in congested traffic*. Journal of Advanced Transportation, 2007. **41**(2): p. 145-170.
103. Chung, K., J. Rudjanakanoknad, and M.J. Cassidy, *Relation between traffic density and capacity drop at three freeway bottlenecks*. Transportation Research Part B: Methodological, 2007. **41**(1): p. 82-95.
104. Cassidy, M.J. and J. Rudjanakanoknad, *Increasing the capacity of an isolated merge by metering its on-ramp*. Transportation Research Part B: Methodological, 2005. **39**(10): p. 896-913.

105. Bertini, R.L. and M.T. Leal, *Empirical study of traffic features at a freeway lane drop*. Journal of Transportation Engineering, 2005. **131**(6): p. 397-407.
106. Srivastava, A. and N. Geroliminis, *Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model*. Transportation Research Part C: Emerging Technologies, 2013. **30**: p. 161-177.
107. Oh, S. and H. Yeo, *Impact of stop-and-go waves and lane changes on discharge rate in recovery flow*. Transportation Research Part B: Methodological, 2015. **77**: p. 88-102.
108. Leclercq, L., et al. *Capacity drops at merges : analytical expressions for multilane freeways*. in *Transportation Research Board 95th Annual Meeting*. 2016.
109. Lighthill, M.J. and G.B. Whitham, *On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads*. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 1955.
110. Richards, P.I., *Shock Waves on the Highway*. Operations Research, 1956.
111. Amer, A., et al., *Traffic Incident Management Gap Analysis Primer*. 2015.
112. Neudorff, L.G., et al., *Freeway management and operations handbook*. 2003.
113. Carson, J.L., *Best practices in traffic incident management*. 2009.
114. Agrawal, R., T. Imieliński, and A. Swami. *Mining association rules between sets of items in large databases*. in *Acm sigmod record*. 1993. ACM.
115. Hahsler, M., B. Grün, and K. Hornik, *A computational environment for mining association rules and frequent item sets*. 2005.
116. Pettet, G., et al. *Incident Analysis and Prediction Using Clustering And Bayesian Network*. in *2017 IEEE International Conference on Smart City Innovations*. 2017.
117. Bujang, M.A., N. Sa'at, and T.M.I.T.A. Bakar, *Determination of minimum sample size requirement for multiple linear regression and analysis of covariance based on experimental and non-experimental studies*. Epidemiology Biostatistics and Public Health, 2017. **14**(3): p. e12117-1--e12117-9.
118. Collett, D., *Modelling survival data in medical research*. 2003.
119. Bhat, C.R., *Duration modeling*. In: *Handbook of Transport Modelling*. 2000.
120. Washington, S.P., M.G. Karlaftis, and F. Mannering, *Statistical and econometric methods for transportation data analysis*. 2010: CRC press.
121. Bartlett, J.E., J.W. Kotrlik, and C.C. Higgins, *Organizational research: Determining appropriate sample size in survey research*. Information Technology, Learning, and Performance Journal, 2001.
122. Kelley, K. and S.E. Maxwell, *Sample Size for Multiple Regression: Obtaining Regression Coefficients That Are Accurate, Not Simply Significant*. Psychological Methods, 2003.
123. Hanley, J.a. and E.E.M. Moodie, *Sample size, precision and power calculations: a unified approach*. Biometrics & Biostatistics, 2011.
124. Aggarwal, C.C. and S. Sathe, *Outlier ensembles: An introduction*. 2017.
125. Beaton, A.E. and J.W. Tukey, *The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data*. Technometrics, 1974.
126. Huber, P.J., *Robust Estimation of a Location Parameter*. The Annals of Mathematical Statistics, 1964.
127. Wold, S., K. Esbensen, and P. Geladi, *Principal component analysis*. Chemometrics and Intelligent Laboratory Systems, 1987.

128. Kaufman, L. and P.J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis (Wiley Series in Probability and Statistics)*. 1990.
129. Hinneburg, A., C.C. Aggarwal, and D.a. Keim, *What is the Nearest Neighbor in High Dimensional Spaces?* Proceedings of the 26th VLDB Conference, 2000.
130. Liu, F.T., K.M. Ting, and Z.H. Zhou. *Isolation forest*. in *Proceedings - IEEE International Conference on Data Mining, ICDM*. 2008.
131. Kolenikov, S. and G. Angeles, *The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomic Indices*. Chapel Hill: Carolina Population Center, University of North Carolina., 2004.
132. Chang, G-L., M. Raqib, and E. Igbinosun, *Performance evaluation and benefit analysis for CHART*. Coordinated Highways Action Response Team, Maryland Department of Transportation, College Park, 2017.
133. Elefteriadou, L., R.P. Roess, and W.R. McShane, *Probabilistic nature of breakdown at freeway merge junctions*. Transportation Research Record, 1995(1484).
134. Persaud, B., S. Yagar, and R. Brownlee, *Exploration of the breakdown phenomenon in freeway traffic*. Transportation Research Record: Journal of the Transportation Research Board, 1998(1634): p. 64-69.
135. Yi, H. and T. Mulinazzi, *Observed distribution patterns of on-ramp merge lengths on urban freeways*. Transportation Research Record: Journal of the Transportation Research Board, 2007(2023): p. 120-129.
136. Leclercq, L., et al., *Capacity drops at merges: analytical expressions for multilane freeways*. Transportation Research Record: Journal of the Transportation Research Board, 2016(2560): p. 1-9.
137. Duret, A., J. Bouffier, and C. Buisson, *Onset of congestion from low-speed merging maneuvers within free-flow traffic stream: Analytical solution*. Transportation Research Record, 2010. **2188**(1): p. 96-107.
138. Leclercq, L., J.A. Laval, and N. Chiabaut, *Capacity drops at merges: An endogenous model*. Procedia-Social and Behavioral Sciences, 2011. **17**: p. 12-26.
139. Leclercq, L., et al., *Capacity drops at merges: New analytical investigations*. Transportation Research Part C: Emerging Technologies, 2016. **62**: p. 171-181.
140. Leclercq, L., et al., *Capacity drops at merges: analytical expressions for multilane freeways*. Transportation Research Record, 2016. **2560**(1): p. 1-9.
141. Lighthill, M.J. and G.B. Whitham, *On kinematic waves II. A theory of traffic flow on long crowded roads*. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 1955. **229**(1178): p. 317-345.
142. Richards, P.I., *Shock waves on the highway*. Operations research, 1956. **4**(1): p. 42-51.
143. Daamen, W., M. Loot, and S. Hoogendoorn, *Empirical Analysis of Merging Behavior at Freeway On-Ramp*. Transportation Research Record: Journal of the Transportation Research Board, 2010.
144. Oehlert, G.W., *A note on the delta method*. The American Statistician, 1992. **46**(1): p. 27-29.
145. Laval, J.A. and L. Leclercq, *Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model*. Transportation Research Part B: Methodological, 2008. **42**(6): p. 511-522.

146. Munjal, P.K., Y.-S. Hsu, and R.L. Lawrence, *Analysis and validation of lane-drop effects on multi-lane freeways*. Transportation Research/UK/, 1971.
147. Elefteriadou, L.A., *The highway capacity manual 6th edition: A guide for multimodal mobility analysis*. ITE Journal, 2016. **86**(4).
148. Hallenbeck, M.E., O.I. Selezneva, and R. Quinley, *Verification, refinement, and applicability of long-term pavement performance vehicle classification rules*. 2014, United States. Federal Highway Administration. Office of Infrastructure
149. Zhou, J., et al., *Estimating Passenger Car Equivalents on Level Freeway Segments Experiencing High Truck Percentages and Differential Average Speeds*. Transportation Research Record, 2018. **2672**(15): p. 44-54.
150. Hancock, M.W. and B. Wright, *A Policy on Geometric Design of Highways and Streets*. 2013, AASHTO.
151. Lee, S.E., E.C. Olsen, and W.W. Wierwille, *A comprehensive examination of naturalistic lane-changes*. 2004, United States. National Highway Traffic Safety Administration.
152. Ping, Y., T. Zongzhong, and Z. Qiang, *Consistency of input-output model and shockwave analysis in queue and delay estimations*. Journal of Transportation Systems Engineering and Information Technology, 2008. **8**(6): p. 146-152.
153. Rakha, H. and W. Zhang. *Consistency of shock-wave and queuing theory procedures for analysis of roadway bottlenecks*. in *84th Annual Meeting of the Transportation Research Board, Washington, DC*. 2005. Citeseer.
154. Park, S., *Consistency of Queueing Analysis and Shock Wave Analysis*. Journal of Transport Research, 2018. **25**(1): p. 1-13.
155. Taylor, N.B., et al., *Modelling delay saving through pro-active incident management techniques*. European transport research review, 2017. **9**(4): p. 48.