TRAFFIC DELAY ESTIMATION, OPERATIONAL STRATEGY ANALYSIS AND RISK ASSESSMENT FOR WORK ZONES

Weng Jinxian

NATIONAL UNIVERSITY OF SINGAPORE
2011
TRAFFIC DELAY ESTIMATION, OPERATIONAL STRATEGY ANALYSIS AND RISK ASSESSMENT FOR WORK ZONES

Weng Jinxian

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF CIVIL & ENVIRONMENTAL ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE
2011
To my wife Bao Nahua and my parents

You all deserve the pride!
ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest appreciation and sincere gratitude to my supervisors, Associate Professor Meng Qiang and Professor Fwa Tien Fang, for their patient guidance and countless invaluable suggestions throughout my entire PhD candidature. Without encouragement and generous support given by Professor Meng, I would not be able to start my PhD study. He not only supervises me on research methodologies but also teaches me the principles of being an upright man. His influence on me will be priceless to my future career and life. I am indebted to Professor Fwa for his motivating words and constructive suggestions throughout my research work.

I am grateful to the National University of Singapore for granting the research scholarship during the course of my doctoral degree. I owe special thanks to Mr. Foo Chee Kiong, Madam Yap-Chong Wei Leng, Madam Theresa Yu-Ng Chin Hoe for their assistance in handling the tools and software required for my research work. Their kind cooperation has allowed me to complete my research smoothly.

I wish to thank my colleagues and friends Liu Zhiyuan, Wang Tingsong, Qu Xiaobo and Wang Xinchang for all kinds of support and encouragement.

Finally, I want to dedicate this accomplishment to my wife and my parents, for their endless love, support and encouragement.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ...............................................................................................i

TABLE OF CONTENTS .............................................................................................. iii

SUMMARY .................................................................................................................. vii

LIST OF TABLES ......................................................................................................... ix

LIST OF FIGURES ....................................................................................................... xi

LIST OF ABBREVIATIONS ....................................................................................... xiii

LIST OF SYMBOLS ..................................................................................................... xv

## CHAPTER 1 INTRODUCTION .............................................................................. 1

1.1 Preamble ............................................................................................................ 1

1.2 Background ....................................................................................................... 2

1.2.1 Work Zone Operational Strategy ................................................................. 2

1.2.2 Traffic Delay at Work Zones ...................................................................... 4

1.2.3 Vehicle Crash Risk at Work Zones ............................................................... 5

1.3 Research Objectives ......................................................................................... 7

1.4 Research Scope .............................................................................................. 8

1.5 Organization of Thesis .................................................................................... 9

## CHAPTER 2 LITERATURE REVIEW .............................................................. 15

2.1 Capacity Estimation at Work Zones ............................................................... 15

2.1.1 Work Zone Capacity Definition .................................................................. 15

2.1.2 Factors Affecting Work Zone Capacity ....................................................... 16

2.1.3 Work Zone Capacity Estimation Methods ............................................... 19

2.2 Models and Tools Applicable to Traffic Delay Estimation ......................... 22

2.2.1 Macroscopic Analytical Models .................................................................. 23

2.2.2 Macroscopic Simulation Models .................................................................. 26

2.2.3 Microscopic Simulation Tools ..................................................................... 27

2.2.4 Cellular Automata Models ......................................................................... 29

2.3 Work Zone Operational Strategy ................................................................. 32

2.3.1 Single Work Zone Operational Strategy ................................................... 32

2.3.2 Subwork Zone Operational Strategy ......................................................... 33

2.4 Vehicle Crash Risk Assessment at Work Zones ........................................... 34

2.4.1 Collective Approach: Vehicle Crash Frequency Modeling at Work Zones ........................................................................................................... 35

2.4.2 Individual Approach: Occurrence Probability of Vehicle Crashes at Work Zones........................................................................................................... 37

2.5 Identified Research Limitations and Gaps ....................................................... 40
9.1 Outcomes and Contributions.................................................................................................202
  9.1.1 Work Zone Traffic Delay Estimation.............................................................................203
  9.1.2 Optimal Subwork Zone Operational Strategy..........................................................205
  9.1.3 Risk Assessment at Work Zones ...............................................................................207
9.2 Recommendations for Future Work................................................................................209

REFERENCES........................................................................................................................................212

APPENDIX: CURRICULUM VITAE ..........................................................................................228
SUMMARY

Work zone projects are necessary for maintaining a good level of service for a road system. In addition to the work zone costs for contractors of completing a work zone project, a work zone project also results in long travel delays and high accident risk for road users. It is thus important to estimate traffic delay, analyze the work zone operational strategy and assess the accident risk for work zone projects. This thesis aims to develop the proper methodologies/models to deal with these three practical issues.

The thesis first builds a novel heterogeneous cellular automata (HCA) model to estimate work zone traffic delays. Compared to the existing cellular automata models, the HCA model possesses three unique characteristics — improved forwarding rules to characterize longitudinal vehicular movements, improved lane changing rules to describe lateral vehicular movements, and randomization probability functions. A case study shows that the HCA model is able to estimate work zone traffic delay more precisely than the other models. This thesis then proposes a decision tree-based model to estimate work zone capacity, which is a key factor in measuring traffic delays at work zones. The statistical comparison results demonstrate that the proposed decision tree-based work zone capacity estimation model outperforms the existing work zone capacity models.

To analyze the work zone operational strategy, this thesis develops two minimization models and solution algorithms for the optimal subwork zone
operational strategy. The first model comes from the system optimum viewpoint, with the objective of minimizing total work zone cost by taking into account variable traffic speeds and the constraint that each subwork zone has uniform length. A genetic simulated annealing method is employed to solve this model. The second model aims to minimize the maintenance costs concerned by the contactors, subject to the queue length and travel delay constraints. A trial-and-error solution method with practical implications is designed. Numerical examples are employed to assess the proposed minimization models and the corresponding solution methods.

The last part of this thesis is concerned with the work zone risk assessment. It first presents a new approach for estimating the rear-end crash risk at work zones by using real work zone traffic data. A probabilistic quantitative risk assessment (QRA) model is subsequently put forward to quantitatively estimate the vehicle occupant’s casualty risk. This model consists of a crash frequency estimation function, an event tree and a consequence estimation model, in which some input parameters are assumed to be random variables to reflect their uncertainty.

This thesis provides a useful guideline which traffic engineers can use to estimate work zone delays and capacities, find an optimal subwork zone operational strategy and assess work zone risk.
LIST OF TABLES

Table 3.1 Details of different types of vehicles ....................................................49
Table 3.2 Calibration results for the HCA model .................................................56
Table 3.3 Calibrated randomization probabilities in expressway work zones......59
Table 3.4 Calibrated randomization probabilities in arterial work zones .........60
Table 3.5 Estimated coefficients associated with the variables for the
randomization probability functions.............................................................60
Table 3.6 Error tests for the individual vehicle speed...........................................67
Table 3.7 Statistical simulated traffic delay results from the proposed HCA model
and PARAMICS............................................................................................72
Table 3.8 Execution times for the HCA model and PARAMICS .........................72
Table 4.1 Collected work zone datasets for training, checking and evaluating
decision tree-based capacity estimation model.............................................90
Table 4.2 Extracted “If-then” rules from the optimal decision tree......................93
Table 4.3 Values for 18 work zone scenarios for evaluating the decision
tree-based model ...........................................................................................94
Table 4.4 Comparison of estimated work zone capacities from the decision
tree-based model with four existing capacity estimation models .................97
Table 5.1 Model parameters used in the numerical example..............................124
Table 5.2 Time dependent traffic flow and average travel speed......................125
Table 5.3 Optimal subwork zone operational strategy with a given number of
subwork zones at medium traffic scenario..................................................127
Table 5. 4 Optimal subwork zone operational strategy from the minimization
model (5.25)-(5.31) using different user delay estimation approach............128
Table 6.1 Input parameters for total maintenance cost minimization model.....143
Table 6.2 Average hourly traffic flow and average travel speed......................144
Table 6.3 Maximum queue length at project start time of y1=9:20 am ..........147
Table 6.4 Comparisons of results from the total work zone cost model and the
total maintenance cost model......................................................................149
Table 7.1 Truncated normal distribution parameters of MADR at dry pavement
conditions....................................................................................................157
Table 7.2 Statistical comparisons of rear-end crash risk for different vehicle types and lane positions at the expressway work zone (sample size=6548)          162
Table 7.3 Statistical results of rear-end crash risk for different vehicle types and lane positions at the arterial work zone (sample size=2802)          163
Table 7.4 Statistical results of rear-end crash risk (R) for different vehicle types and lane positions at the expressway and arterial work zones (sample size=9350)  163
Table 7.5 Stepwise selection results of the explanatory variables for Model 1 165
Table 7.6 Statistical results of the rear-end crash risk model for the expressway work zone (sample size=36) 166
Table 7.7 Statistical results of the rear-end crash risk model for the arterial work zone (sample size=32) 168
Table 7.8 Summary of the combined rear-end crash risk model for the expressway and arterial work zones (sample size=68) 170
Table 8.1 Ohio work zone crash accident data in 2002 176
Table 8.2 Trial-and-error results for work zone crash frequency 177
Table 8.3 Input parameters for the numerical example 191
Table 8.4 Estimated probabilities of intermediate events in the event tree 192
LIST OF FIGURES

Figure 1.1 Structure of the thesis ................................................................. 10
Figure 2.1 Delay estimated by deterministic queuing model ...................... 24
Figure 3.1 Schematic representation of work zone in simulation .................. 48
Figure 3.2 Comparison of longitudinal vehicle trajectories in Lane 1 and 2 from HCA model against the field data .............................................. 63
Figure 3.3 Comparison of lateral vehicle trajectories in Lane 1 from the HCA model against the field data ............................................................... 64
Figure 3.4 Comparison of lane change duration from the HCA model against the field data ............................................................................. 65
Figure 3.5 Comparison of average travel time from the HCA simulation versus field data .......................................................... 69
Figure 3.6 Comparisons of traffic delay from the HCA simulation against the field data .............................................................. 71
Figure 3.7 Effects of the transition area length and activity area length on $D_{wz}$ .. 74
Figure 3.8 Effects of the transition area length and activity area length on $d_{wz}$ ... 75
Figure 3.9 Impacts of traffic flow and heavy vehicle percentage on $D_{wz}$ ........ 76
Figure 4.1 Three steps for the development of a decision tree ..................... 87
Figure 4.2 Learning process of the decision tree model for work zone capacity estimation ............................................................................. 91
Figure 4.3 The optimal decision tree for work zone capacity estimation ......... 92
Figure 4.4 Comparison of estimated and observed work zone capacity .......... 99
Figure 5.1 Two roadway pictures shot in Singapore .................................... 101
Figure 5.2 Two work zone operational strategies ....................................... 103
Figure 5.3 Illustrations of three cases in the queuing delay estimation .......... 109
Figure 5.4 Three components of total work zone cost incurred by the optimal subwork zone operational strategy with a given number of subwork zones ..... 130
Figure 5.5 Impacts of parameters $d_1$ and $d_2$ on the optimal subwork zone operational strategy ......................................................... 131
Figure 5.6 The number of subwork zones in the optimal subwork zone operational strategy with respect to a fixed work zone setup time ............ 133
Figure 5.7 Three components of the total work zone cost with respect to a given number of subwork zones for the heavy traffic scenario when $d_i=0.1h/zone$.

Figure 5.8 Three components of the total work zone cost with respect to a given number of subwork zones for the medium traffic scenario when $d_i=0.5h/zone$.

Figure 5.9 Three components of the total work zone cost with respect to a given number of subwork zones for the light traffic scenario when $d_i=0.5h/zone$.

Figure 6.1 Total maintenance cost versus the number of subwork zones.

Figure 6.2 Average travel delay versus project start time ($y_1$).

Figure 6.3 Maximum queue length versus project start time ($y_1$).

Figure 7.1 Work zone sites for field data collection.

Figure 7.2 Estimated $DRAC$ versus lane traffic flow, heavy vehicle percentage, lane position and work zone type.

Figure 8.1 A flowchart of QRA model formulation.

Figure 8.2 Structure of event tree for work zone crash.

Figure 8.3 Probability density functions for the probabilities of intermediate events in the event tree.

Figure 8.4 Analysis of individual risk in work zone.

Figure 8.5 Analysis of F/N curve for casualties in work zone.

Figure 8.6 Impact of uncertainty associated with the probability of intermediate event.

Figure 8.7 Influence of $ERT$ on individual risk.

Figure 8.8 Influence of speed on individual risk.
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Age</td>
</tr>
<tr>
<td>AASHTO</td>
<td>American Association of State Highway and Transportation Officials</td>
</tr>
<tr>
<td>ADT</td>
<td>Average Daily Traffic</td>
</tr>
<tr>
<td>AHT</td>
<td>Average Hourly Traffic</td>
</tr>
<tr>
<td>AL</td>
<td>Alcohol</td>
</tr>
<tr>
<td>BL</td>
<td>Brake Light</td>
</tr>
<tr>
<td>CA</td>
<td>Cellular Automata</td>
</tr>
<tr>
<td>CBD</td>
<td>Central Business District</td>
</tr>
<tr>
<td>CT</td>
<td>Crash Type</td>
</tr>
<tr>
<td>CU</td>
<td>Crash Unit</td>
</tr>
<tr>
<td>DQM</td>
<td>Deterministic Queuing Model</td>
</tr>
<tr>
<td>DRAC</td>
<td>Deceleration Rate to Avoid the Crash</td>
</tr>
<tr>
<td>ERT</td>
<td>Emergency Response Time</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information System</td>
</tr>
<tr>
<td>GSA</td>
<td>Genetic Simulated Annealing</td>
</tr>
<tr>
<td>HCA</td>
<td>Heterogeneous Cellular Automata</td>
</tr>
<tr>
<td>HCM</td>
<td>Highway Capacity Manual</td>
</tr>
<tr>
<td>LC</td>
<td>Light Condition</td>
</tr>
<tr>
<td>LTA</td>
<td>Land Transport Authority</td>
</tr>
<tr>
<td>MADR</td>
<td>Maximum Available Deceleration Rate</td>
</tr>
<tr>
<td>MPE</td>
<td>Mean Percent Error</td>
</tr>
<tr>
<td>MSHA</td>
<td>Maryland State Highway Administration</td>
</tr>
<tr>
<td>MUTCD</td>
<td>Manual on Uniform Traffic Control Devices</td>
</tr>
<tr>
<td>PCE</td>
<td>Passenger Car Equivalent</td>
</tr>
<tr>
<td>PDO</td>
<td>Property Damage Only</td>
</tr>
<tr>
<td>PET</td>
<td>Post Encroachment Time</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>QRA</td>
<td>Quantitative Risk Assessment</td>
</tr>
<tr>
<td>QUEWZ</td>
<td>Queue and User Cost Evaluation of Work Zones</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>RMSPE</td>
<td>Root Mean Square Percent Error</td>
</tr>
<tr>
<td>S</td>
<td>Severity</td>
</tr>
<tr>
<td>SA</td>
<td>Simulated Annealing</td>
</tr>
<tr>
<td>SAS</td>
<td>Statistical Analysis Software</td>
</tr>
<tr>
<td>SSM</td>
<td>Surrogate Safety Measure</td>
</tr>
<tr>
<td>TDC</td>
<td>Traffic Data Collector</td>
</tr>
<tr>
<td>TTC</td>
<td>Time to Collision</td>
</tr>
<tr>
<td>U</td>
<td>Theil’s Inequality Coefficient</td>
</tr>
<tr>
<td>VT</td>
<td>Vehicle Type</td>
</tr>
</tbody>
</table>
LIST OF SYMBOLS

\( a_n(v_n^{-1}, k) \)  
Acceleration rate of vehicle \( n \) which belongs to type \( k \) at time \( t \)

\( b_n^t \)  
Brake light status of vehicle \( n \) at time \( t \)

\( c_1 \)  
Fixed cost of installing and removing devices for alerting and redirecting vehicles

\( c_2 \)  
Variable cost of installing and removing traffic control devices in the workspace

\( c_3 \)  
Fixed cost of traffic control devices for altering and redirecting vehicles

\( c_4 \)  
Variable cost of traffic control devices placed in the workspace

\( c_a \)  
Average cost per accident

\( c_{sh} \)  
Average delay cost per vehicle-hour

\( C_{0h+j} \)  
Freeway capacity during time interval \([u_0+j, u_0+j+1]\)

\( C_{wh+j} \)  
Work zone capacity during time interval \([u_0+j, u_0+j+1]\)

\( d_1 \)  
Fixed setup time for a subwork zone

\( d_2 \)  
Variable work time per kilometer for a subwork zone

\( d'_a \)  
Available front gap between vehicle \( n \) and its leading vehicle

\( d'_{n,\text{eff}} \)  
Effected available front gap

\( d'_{n,n,f} \)  
Available front gap between vehicle \( n \) and its front neighboring vehicle in the target lane

\( d'_{n,b,h} \)  
Available back gap between vehicle \( n \) and its back neighboring vehicle in the target lane
\( d_n(k) \)  Deceleration rate of vehicle \( n \) which belongs to type \( k \)

\( d_{\text{security}} \)  Security longitudinal distance between two vehicles

\( D_{sz} \)  Average traffic delay per work zone

\( DRAC_n^{n-1} \)  The deceleration rate of the vehicle \( n \) to avoid the crash with the vehicle \( n-1 \)

\( E \)  The allowed error range

\( E_f \)  Passenger-car equivalent for the truck

\( f_d \)  The traffic directional distribution factor

\( f_g \)  Adjustment factor for grade

\( f_H \)  Heavy vehicle flow

\( f_L \)  Light vehicle flow

\( f_w \)  Adjustment factor for lane width

\( f_{\text{crash}}(Q,D,L,U) \)  The frequency of work zone crash

\( F_N(x) \)  The frequency that the number of casualties caused by work zone crash, \( x \), is not less than the number \( N \)

\( hv \)  Heavy vehicle percentage

\( IR_F \)  The individual fatality risk

\( IR_I \)  The individual injury risk

\( L_a \)  Length of work zone activity area

\( l_{\text{min}} \)  The minimum length of a subwork zone restricted by traffic engineers

\( l_{\text{max}} \)  The maximum length of a subwork zone

\( l_n(k) \)  Vehicle length of the \( n^{th} \) vehicle
$L_{pos}$  Lane position

$L_t$  Length of work zone transition area

$L_s$  Length of work zone advance warning area

$n_a$  The number of accidents per 100 million vehicle hours

$N_{Fi}$  The number of fatalities in a vehicle of type $i$ involving work zone crash

$N_{Li}$  The number of injuries in a vehicle of type $i$ involving work zone crash

$N_i$  The average number of occupants in a vehicle of type $i$

$N_{run}$  The required number of simulation runs

$N_{FF_r}$  The total number of fatalities in the $r$th accident scenario

$N_{FI_r}$  The total number of injuries in the $r$th accident scenario

$\rho(d_1, d_2, n)$  Total work zone cost reduction ratio for the $n$ number of subwork zones

$\rho(e)$  The uncertainty ratio of upper bound to lower bound for the uncertainty of outcome caused by the uncertain event $e$

$p_1$  Randomization probability outside work zone

$p_2$  Randomization probability in work zone

$p \left(DRAC_{n,j}^{n-1} > MADR_{n,j}\right)$  The probability of a rear-end crash between vehicles $n$ and $n-1$ at time $t$

$p_{n,t}$  The probability of rear-end crash for the vehicle $n$ at time $t$

$p_{seq_r}$  The occurrence probability of the $r$th accident scenario

$P_T$  Proportion of trucks
\( Q_{\text{max}} \) The queue length threshold imposed by land transport authorities

\( \delta \) The length of a uniform time interval in traffic flow distribution

\( \sigma \) The sample standard deviation

\( \Delta T \) The observation time interval

\( t_n^h \) Available time headway of vehicle \( n \) at time \( t \)

\( T_\text{v} \) The travel delay threshold imposed by land transport authorities

\( TL_n \) The number of target lane that the vehicle \( n \) expects to move into

\( T_{\text{wz}} \) Average travel time in work zone

\( T_0 \) Average travel time outside work zone

\( u_0 \) The earliest project start time

\( u_t \) The latest project completion time

\( v_n' \) Speed of vehicle \( n \) at time \( t \), where \( v_n' = 0, 1, 2, \ldots, v_{\text{max}}^k \)

\( v_{\text{max}}^k \) Maximal speed of vehicle type \( k \) (1 for light vehicle; 2 for heavy vehicle)

\( v_{\text{max}}^n \) Maximal speed of vehicle \( n \)

\( v_{\text{max}}^{n,n',h} \) Maximum speed of the back neighboring vehicle corresponding the vehicle \( n \)

\( V_{m+}^{u_0+j} \) Average speed of vehicles passing through work zone during time interval \([u_0+j, u_0+j+1]\)

\( V_{a+}^{u_0+j} \) Average speed of vehicles approaching work zone during time interval \([u_0+j, u_0+j+1]\)

\( V_{\text{pos}} \) Posted speed limit at work zone

\( V_{\text{type}} \) The indicator variable for vehicle types

\( (x, y) \) A subwork zone operational strategy
$(x'_n, y'_n)$ Top left position of vehicle $n$ at time $t$

$W_{type}$ The indicator variable for work zone types

$z_{\alpha/2}$ The threshold value for $100(1 - \alpha)$ percentile confidence interval
For any country, the road system is an important part of its infrastructure which can affect economic development. An efficient road system can promote economic development because it enhances the performance of local transportation. A good level of service for a road system requires the implementation of work zone projects to maintain it. Therefore, various work zone projects, such as pothole patching, crack sealing and pavement resurfacing, are regularly carried out by land transport authorities. For example, the Land Transport Authority of Singapore planned 3,257 work zone projects between May 2009 and March 2010 (LTA, 2009).

Although work zone projects are necessary, they also cause problems. The cost of a work zone project is usually very high. Therefore, finding an optimal operational strategy to minimize work zone costs is a major concern for contractors. Since work zone projects close one or more of the lanes available for traffic, crash risk and traffic delay are the two major issues, which land transport authorities are concerned. In general, the likelihood of vehicle crash in a work zone is higher than that in non-work zones (Ha and Nemeth, 1995; Rouphail et al., 1988; Ullman et al., 2006). Traffic delay occurring in work zones can further lead to many adverse effects, such as lost time, higher fuel consumption and vehicle emission.

For these reasons, contractors have to find an optimal operational strategy and evaluate crash risk and traffic delay before taking effective measures to lower work
zone costs and mitigate the negative impacts.

### 1.2 Background

Many researchers have studied the work zone problems (Martinelli and Xu, 1996; Kim et al., 2001; Qi et al., 2005; Harb et al., 2008; Yang et al., 2008). Their studies can be categorized into three groups. The first focuses on pursuing an optimal work zone operational strategy, including optimal work zone length and best starting time, in order to minimize work zone costs. The second aims to accurately estimate traffic delays at work zones. The third evaluates crash risk and takes corresponding mitigation measures.

#### 1.2.1 Work Zone Operational Strategy

In general, two types of operational strategies are available for the implementation of work zone projects: 1) the single work zone operational strategy, which closes the entire lane in one travel direction during project implementation; and 2) the subwork zone operational strategy which divides the entire work zone into multiple subwork zone segments, with relevant work activities undertaken in one segment at a time (Martinelli and Xu, 1996). Up till now, a variety of models have been developed for the optimization of the single work zone operational strategy (Schonfeld and Chien, 1999; Chien and Schonfeld, 2001; Jiang and Adeli, 2003; Chen
et al., 2005; Yang et al., 2008). However, existing models for single work zone operational strategy are based on the assumption that the length of the entire work zone is unknown. From a practical perspective, such an assumption is incorrect because, in reality, the location and length of the entire work zone are always known by the contractor. Thus, these models cannot be applied to real work zone projects.

Therefore, some researchers (McCoy et al., 1980; Martinelli and Xu, 1996; Chen and Schonfeld, 2006; Tang and Chien, 2010) developed several models to determine the optimal subwork zone operational strategy, assuming a known work zone length. Nonetheless, there remain three limitations. First, total queuing delay may be overestimated because the researchers neglected to account for the fact that a queue may completely disappear before the end of a time interval when the arriving traffic flow is lower than the work zone capacity at that time interval. There is also a flaw in the moving delay estimation formula because researchers do not consider the fact that the vehicle departure rate will be equal to the work zone capacity before the queue completely disappears. Second, the approaching traffic speed is assumed to be constant in these studies. According to the Highway Capacity Manual (HCM, 2000), traffic speed actually varies with traffic volume. Third, there is no uniform length constraint in the existing subwork zone studies. However, each subwork zone usually has a uniform length in reality. Therefore, from a practical viewpoint, there is still a need to find a subwork zone operational strategy considering variable traffic speed, time window and uniform length constraints.

It should also be noted that all previous models for the subwork zone
operational strategy have aimed to minimize total work zone cost, which is the sum of costs for two different groups: road users and contractors. However, contractors are actually more concerned with minimizing their maintenance/construction costs, subject to the queue length and travel delay constraints imposed by the land transport authorities. This is an interesting topic that will be explored in this research.

1.2.2 Traffic Delay at Work Zones

The accurate estimation of traffic delay at work zones is of utmost importance because it is a key step towards taking effective traffic management plans to mitigate the delays. Generally, traffic delay at work zones can be estimated by any one of the following methods: the macroscopic analytical method, the macroscopic simulation method or the microscopic simulation method. The deterministic queuing model, which is one of macroscopic analytical methods, has been used for estimating delays for decades. However, it may lead to an underestimation result because of the neglect of traffic shock-wave delays. Nevertheless, the deterministic queuing model is still widely accepted by traffic engineers because of its simplicity (McCoy et al., 1980; Schonfeld and Chien, 1999; Jiang and Adeli, 2003).

In comparison with the macroscopic traffic simulation methods, the microscopic traffic simulation methods are able to model more detailed traffic dynamics at the individual level and provide relatively high estimation accuracy. Many researchers have developed and applied microscopic simulation models or tools
to estimate traffic delays at work zones (Maze and Kamyab, 1999; Chien et al., 2002; Yang et al., 2008). However, these simulation methods still need to be improved for a number of reasons. First, they require much time and high computational resource to estimate traffic delay. Second, some of these simulation models/tools (e.g., CORSIM) cannot describe the interaction between vehicles and work zone configurations well because the work zone is simulated through a prolonged incident blockage with no transition area. Thus, these types of methods cannot adequately estimate the traffic delays at work zones. Many of the existing microscopic simulation models, such as CELLSIM and TRANSIMS, have been developed based on the NaSch (Nagel and Schreckenberg, 1992) Cellular Automata (CA) model because of its high degree of computational efficiency. Nonetheless, most of these CA-based traffic flow models are only applicable to simulating homogeneous traffic flow. In those CA models for heterogeneous traffic flow, the randomization probability parameter is assigned a hypothetical constant value, resulting in a coarse description of traffic operations.

In this research, an attempt will be made to develop a microscopic model based on the stochastic CA model for simulating heterogeneous work zone traffic. The randomization probability will be related to traffic flow and work zone configuration. The developed heterogeneous CA model will further be applied to estimate traffic delay at work zones.

### 1.2.3 Vehicle Crash Risk at Work Zones

The presence of work zone increases traffic conflicts and causes severe traffic
safety problems. As have been shown by many researchers (Ha and Nemeth, 1995; Rouphail et al., 1988; Ullman et al., 2006), the occurrence likelihood of a severe crash in work zone is higher than that in non-work zone and severe crashes lead to casualties. Some researchers have further indicated that the activity area is the predominant location of work zone crashes, regardless of road type, and that rear-end crashes are the predominant type of crash (Pigman and Agent, 1990; Garber and Zhao, 2002).

Numerous studies have been conducted on the analysis of rear-end crashes (e.g., Abdel-Aty and Abdelwahab, 2004; Qi et al., 2005; Kim et al., 2007; Harb et al., 2008). In these studies, historical accident data are utilized to identify the causal factors and injury severity using statistical techniques. This is an efficient method of analyzing vehicle crash risk. However, there is a possibility that no historical accident data will be available, especially for a newly-built road or a newly-proposed work zone. In order to assess vehicle crash risk at work zones for which historical accident data are not available, an alternative method that avoids making use of historical accident data must be explored.

Although various models have been developed for crash risk assessment in previous studies, most of these models emphasize estimating the occurrence likelihood or frequency of a work zone crash. However, in practice, traffic safety engineers seem to pay more attention to the casualty risk, e.g., the likelihood of a driver/passenger being killed or injured in a work zone and the relationship between the frequency and consequence of a work zone crash. Without simultaneously taking into account the frequency and consequence of work zone crash, such work zone risk assessment models cannot be applied to assess the overall safety level of work zone. Therefore, an effective methodology needs to be proposed to evaluate the overall
safety level of work zone by simultaneously taking into account the frequency and consequence of work zone crashes.

1.3 Research Objectives

Although a number of models and methodologies have been proposed to estimate traffic delay, analyze the optimal subwork zone operational strategy and assess the risks of work zones, the following limitations and shortcomings still exist:

- There are some flaws in user delay estimation in past studies for the optimal subwork zone operational strategy problem. In addition, these studies adopted unrealistic constant traffic speeds and do not take into account the uniform subwork zone length constraint.

- The current microscopic simulation models and tools take a significant amount of time and computational resources to estimate traffic delays. Although many CA-based models have high computational efficiency, they are applicable only to homogeneous traffic and are too coarse to properly represent the complex driver behavior at work zones.

- A majority of existing crash risk assessment models were developed based on the historical data. They cannot assess vehicle crash risk at work zones where historical accident data are unavailable. In addition, the current crash risk assessment models cannot assess the casualty risk at work zones, which traffic safety engineers are most concerned with.

The objective of this research is therefore to develop models for traffic delay...
estimation, operational strategy analysis and risk assessment for work zones. More specifically, the following research tasks have been conducted to achieve the objective:

1) Develop a heterogeneous CA model to estimate traffic delay at work zones and analyze the impacts of work zone configuration on traffic delay.

2) Mitigate the shortcomings of existing models and methodologies used for the optimal subwork zone operational strategy problem.

3) Propose an effective methodology to evaluate vehicle crash risk at work zones where the historical accident data are not available.

4) Evaluate the vehicle occupant’s casualty risk, by combining the occurrence frequency and consequence of work zone vehicle crashes.

1.4 Research Scope

To accurately estimate traffic delay at work zones, the randomization probability parameter of the CA model should be formulated as a function of the activity area length, the transition area length and the volumes of different types of vehicles traveling through the work zone. For simplicity, one assumption is that the heterogeneous work zone traffic consists of two different vehicle types: light and heavy. The established heterogeneous CA model is calibrated and validated using data from work zones located on the six-lane two-way expressways and arterial roads of Singapore.
To determine the optimal subwork zone operational strategy, a total work zone cost minimization model and a total maintenance cost minimization model are developed from the systemic viewpoint and from the contractor’s perspective, respectively. Due to its simplicity, the deterministic queuing model is used to estimate the total user delay associated with a subwork zone operational strategy. Since the heterogeneous CA model is capable of capturing vehicle interactions at work zones in detail, which may provide highly accurate estimates of traffic delays, the model is also applied for this purpose.

When no historical accident data are available to make a risk assessment, work zone traffic data can be utilized to assess the vehicle crash risk. Hence, this research proposes a methodology based on work zone traffic data. Traffic data at work zone activity areas are collected to test the methodology. To evaluate the casualty risk at work zones, the research uses a probabilistic quantitative risk assessment (QRA) model which simultaneously takes into account the frequency and consequence of crashes occurring in work zones.

1.5 Organization of Thesis

This thesis is organized into nine chapters illustrating the steps taken to achieve the objectives of the research. The flowchart of the thesis is shown in Figure 1.1. Its structure is briefly discussed below.
Figure 1.1 Structure of the thesis

Chapter 1 introduces the background of the study. The importance and need for this research are also discussed. In addition, the objectives and scope of the research are highlighted.

Chapter 2 provides a critical review of past studies related to traffic delay.
estimation, operational strategy analysis and risk assessment at work zones. It first defines work zone capacity and the existing methodologies and models for estimating work zone capacity. Subsequently, a detailed discussion of the available models and tools for traffic delay estimation at work zones is presented, followed by a discussion of past studies on work zone operational strategy. The collective and individual approaches to the assessment of crash risk at work zones are also discussed. Finally, based on the literature review, the potential gaps and limitations are identified.

Chapter 3 presents a heterogeneous cellular automata (HCA) model for estimating traffic delay at work zones. The proposed HCA model includes the improved forwarding rules to update longitudinal speeds and positions of work zone vehicles. The randomization probability parameter is formulated as a function of the activity area length, the transition area length and the volumes of different types of vehicles traveling through the work zone. In addition, the HCA model incorporates a new and realistic lateral speed and position updating rule so that the simulation of vehicle’s lateral movement in work zone is close to the reality. The HCA model is then calibrated and validated microscopically and macroscopically using the real work zone data. Finally, the HCA model is applied to estimate traffic delay at work zones.

Chapter 4 aims to develop a decision tree-based model to accurately estimate freeway work zone capacity, taking into account sixteen influencing factors. The $F$-test splitting criterion is employed to split nodes to grow the tree. A post-pruning approach, which aims to minimize the mean squared prediction error of the checking data, is employed to prune the grown decision tree. Freeway work zone capacity data,
collected from fourteen states and cities, are used to train, check and evaluate the
decision tree-based capacity estimation model. Finally, the estimation accuracy of the
decision tree-based model is compared with that of the existing work zone capacity
models.

Chapter 5 focuses on determining an optimal subwork zone strategy for the
short-term work zone projects in four-lane two-way freeways with time window and
uniform subwork zone length constraints. The deterministic queuing model and the
HCA model are both employed to estimate total user delay. After the estimation of the
user delay, this chapter proceeds to build a total work zone cost minimization model
subject to time window and uniform length constraints, from the systemic perspective.
It also presents a variation of the minimization model which examines the impact of
unequal subwork zone length constraint. Since these models belong to the
mixed-integer non-differentiable optimization problems, an iterative algorithm
embedding with the genetic simulated annealing (GSA) method is proposed for
solving them. Finally, a numerical example is used to investigate the effectiveness of
the proposed models.

From the contractor’s standpoint, Chapter 6 proposes a total maintenance cost
minimization model to determine the optimal subwork zone operational strategy,
subject to queue length and travel delay constraints. An efficient trial-and-error
method is designed to solve the proposed model. A numerical example is used to test
the model and solution method. Finally, a model comparison is carried out between
the total maintenance cost minimization model and the total work zone cost
Chapter 7 proposes a methodology for assessing the rear-end crash risk at work zone activity areas and analyzes the impacts of contributing factors using work zone traffic data when the historical accident data are not available. The deceleration rate to avoid the crash (DRAC) is used to measure the rear-end crash risk. Based on the arterial and expressway work zone traffic data in Singapore, three rear-end crash risk models are developed to examine the relationship between rear-end crash risk and its contributing factors.

Chapter 8 develops a probabilistic QRA model to evaluate the casualty risk combining frequencies and consequences of all accident scenarios triggered by work zone crashes. The casualty risk is measured by individual risk and societal risk. The proposed probabilistic QRA model consists of the estimation of work zone crash frequency, an event tree and consequence estimation models. There are seven intermediate events in the event tree — age, crash unit, vehicle type, alcohol, light condition, crash type and severity. Since the estimated probabilities of some intermediate events may have large uncertainty, this uncertainty can be characterized by a random variable. The consequence estimation model takes into account the combination effects of speed and emergency medical service response time on the consequence of work zone crash. Finally, a numerical example, based on the Southeast Michigan work zone crash accident data, is used to test the proposed QRA model.

Finally, Chapter 9 summarizes the main findings drawn from the current
research and highlights their contribution to the state-of-the-art. It also provides directions and recommendations for future research.
CHAPTER 2 LITERATURE REVIEW

This chapter presents a critical review of the existing studies on traffic delay estimation, operational strategy analysis and risk assessment for work zones. Since work zone capacity is an important factor affecting traffic delay estimation, this chapter first gives the definition of work zone capacity and describes the existing models used to estimate work zone capacity. Then, a detailed discussion of the models and tools available for traffic delay estimation at work zones is presented, followed by discussions of past studies of work zone operational strategy analysis. After this, the collective and individual approaches to the assessment of crash risk at work zones are reviewed. Arising from the literature review, the potential gaps and limitations are identified finally.

2.1 Capacity Estimation at Work Zones

2.1.1 Work Zone Capacity Definition

There is still no unified definition of work zone capacity; a few of the existing definitions are as follows:

- The flow rate derived from three-minute time intervals during congested conditions (California) (Kermode et al., 1970);
- The hourly traffic volume under congested traffic conditions (Texas) (Dudek and Richards, 1981);
• The flow rate at which traffic conditions quickly change from uncongested to queue conditions (North Carolina) (Dixon et al., 1996);
• The maximum recorded five-minute flow rate (Pennsylvania) (Jiang, 1999);
• The flow rate just before a sharp speed drop (Indiana) (Jiang, 1999);
• The highest flow sustained during a 15-minute time period that is either before a rapid speed drop or after a rapid speed increase (Illinois) (Benekohal et al., 2003).

2.1.2 Factors Affecting Work Zone Capacity

There are many factors that could affect work zone capacity. Some have greater significance than others, and some are combined and measured by their joint influence within capacity analysis. Most factors can be grouped into one of the following five major categories.

(1) Work zone configurations

• *Number of closed lanes and number of open lanes.* Measurements made at freeway work zones in Texas (Dudek and Richards 1981; Krammes and Lopez, 1994) and North Carolina freeways (Dixon et al., 1997) show clearly that the work zone capacity varies significantly with the number of freeway lanes as well as the number of lane closures.

• *Lane closure location (Left/Right).*

• *Work zone grade.* Kim et al. (2001) found that the presence of grades may exacerbate the flow constriction in work zones particularly in the presence of
heavy vehicles.

- **Work zone length.** Whether work zone length affects work zone capacity or not is a controversial topic. Kim et al. (2001) claimed that longer length would reduce work zone capacity, while Healsip et al. (2009) found that work zone capacity would not significantly vary with work zone length.

- **Work zone speed.** A lower work zone speed could improve safety, while decreasing the work zone capacity. (Adeli and Jiang, 2003).

**2) Roadway conditions**

- **Road type (Rural/Urban).** Dixon et al. (1996) found that work zone capacity on an urban road is usually 20%–30% higher than that on a rural road.

- **Ramp.** Ramps’ proximity to the work zone, especially the entrance ramps inside the work zone activity area, can create traffic turbulence, resulting in a reduction of the work zone capacity (HCM, 2000).

- **Lane width.** The width of the lanes and the distance to lateral obstructions will both affect capacity. The HCM (2000) suggests a reduction factor of up to 14% to account for the effect of lane width on work zone capacity.

**3) Work activity**

- **Work intensity.** Work zone capacity may decrease as the work intensity increases from lightest (e.g., guardrail installation) to heaviest (e.g., bridge repair). Work zone intensity is classified into three levels (low, medium, or high) in Karim and Adeli (2003). The HCM (2000) suggests a modification of the base capacity value of the work zone to account for the intensity of the
work activity, without actually providing any modification factors or guidelines.

- **Work time (Day/Night).** According to the work of Al-Kaisy and Hall (2001), night construction or maintenance could decrease the work zone capacity because of the reduced traveler’s attention.

- **Work zone duration (Short-term/Long-term).** In general, the average capacity at long-term work zones is greater than that at short-term work zones because commuters and frequent travelers become familiar with the configuration of the long-term work zone.

(4) **Environmental conditions**

- **Weather conditions.** Weather (e.g., snowy, rainy, sunny) usually has a significant impact on work zone capacity. The HCM (2000) suggests 10–20% capacity reductions due to bad weather conditions, without providing any specific guidelines.

(5) **Others**

- **Heavy vehicle percentage.** Since heavy vehicles such as trucks occupy more space on the roadway and move more slowly than passenger cars, a high percentage of heavy vehicles tends to reduce the work zone capacity. Krammes and Lopez (1994) carried out a work zone capacity study and concluded that a high percentage of heavy vehicles has a significant impact on the work zone capacity.

- **Driver composition (Commuters/Non-commuters).** In all situations,
commuters and regular travelers are more familiar with the work zone configuration and traffic control plans than the non-commuters (e.g., tourists). Therefore, the presence of non-commuters could reduce work zone capacity.

- State/city.

### 2.1.3 Work Zone Capacity Estimation Methods

A number of models and methodologies have been developed for estimating work zone capacity.

Firstly, looking at studies on short-term work zones, Krammes and Lopez (1994) developed a capacity estimation model for short-term freeway work zones based on data collected in 33 work zones in Texas between 1987 and 1991. The overall average capacity of 1,600 passenger cars per hour per lane (pcphpl) was used as the base capacity. Adjustment factors included heavy vehicles, the intensity of work activity, the presence of ramps and the number of lanes open through work zone. It was found that the capacities of individual work zones fell within a range of ±10 percent of 1,600 pcphpl.

Kim et al. (2001) presented an interesting method to estimate work zone capacity. They collected traffic data in 12 work zone sites with lane closures on four normal lanes in one direction, and investigated various independent factors contributing to the work zone capacity reduction. A multiple regression model was developed to estimate the capacity by establishing its functional relationship with...
several key independent factors, such as the number of closed lanes, the proportion of heavy vehicles, the grade and the intensity of work activity.

Benekohal et al. (2004) gave a step-by-step approach to estimating work zone capacity for a two-to-one lane closure configuration. They established the relationship between work zone capacity and operating speed, based on extensive data collected from 11 work zones in Illinois. The operating speed in the work zone was expressed as a function of work intensity, lane width, lateral clearance and other factors.

Ping and Zhu (2006) used CORSIM to derive short-term work zone capacities under various network configurations. The parameters tested in their experimental design were the number of open lanes, free-flow speeds along a normal freeway segment and in work zone, the grade, the heavy vehicle percentage, the location of warning signs and the location of closed lanes. The derived capacity was found to range between 1,320 vehicles per hour per lane (vphpl) and 1,920 vphpl, depending on the level of each parameter.

Sarasua et al. (2006) developed a model to estimate the capacities for two-to-one, three-to-two and three-to-one lane closure configurations of interstate highway work zones. In their study, the base capacity depended on the lane closure configuration and the passenger car equivalent (PCE) value varied with traffic speed.

Among studies on long-term work zones, Al-Kaisy et al. (2000) investigated freeway capacity in work zones involving long-term lane closures in terms of the variations in capacity under the effect of important control and extraneous variables in Ontario, Canada. The variables mainly included temporal variations, the grade, the
day of week, and weather conditions. The results showed significant variations in freeway work zone capacity. These variables all exhibited significant effects on capacity.

Al-Kaisy and Hall (2003) examined the capacities of six reconstruction sites in Ontario and then developed a generic multiplicative model to estimate the long-term work zone capacity. The effects of heavy vehicles, driver population, weather, lane configuration, work intensity and light conditions on capacity were each examined and the results suggested that heavy vehicles and driver population were the two most significant factors affecting work zone capacity.

In another study, conducted by Heaslip et al. (2009), analytical models and procedures were proposed to estimate long-term work zone capacity, taking geometric, traffic and work zone related factors into consideration. The analytical models were calibrated using simulation data from CORSIM and validated using field data.

The literature review shows each of the existing models is applicable either to short-term or long-term work zones but not to both. Since long-term work zones generally have higher capacities than short-term work zones, short-term work zone models developed based on the short-term work zone data may greatly underestimate long-term work zone capacity. Similarly, long-term models may overestimate short-term work zone capacity.

Since a large number of factors can affect work zone capacity, it would be inappropriate to estimate capacity using a simple model. The current HCM (2000) provides two distinct work zone capacity estimation guidelines, applying to short-term
work zones and long-term work zones, respectively. However, the HCM requires traffic engineers to decide on the adjustment factors themselves, which may lead to significant estimation errors due to possibly subjective judgments being made. Similar to the existing capacity estimation models, the HCM also excludes some important factors, which may lead to less accurate estimates.

Adeli and Jiang (2003) developed a neural-fuzzy model to estimate capacity. The neural-fuzzy model introduces a neural network of interacting variables that are quantified using a fuzzy inference method. This model accounts for seventeen different variables affecting work zone capacity. Compared with the existing capacity estimation models and guidelines, the neural-fuzzy model is more comprehensive because it incorporates much more important factors and exhibits slightly higher accuracy. However, this model is so complex that it has poor applicability.

2.2 Models and Tools Applicable to Traffic Delay Estimation

Traffic delay is defined as the difference between travel time on a road segment under free-flow conditions and the longer travel time that actually occurs due to some given reasons. This research is concerned with the traffic delays at work zones caused by lane reduction.

In general, traffic delay at work zones can be estimated by any one of the following models and tools — the macroscopic analytical model, the macroscopic simulation model, the microscopic simulation tool or the Cellular Automata model.
This section includes a brief discussion of these models in relation to traffic delay estimation at work zones.

2.2.1 Macroscopic Analytical Models

Deterministic queuing models are the common macroscopic analytical models that are most commonly used for estimating traffic delay and have been used for decades. They are often illustrated using the diagram shown in Figure 2.1. The critical inputs are the demand volume, the freeway capacity, the work zone capacity, and the work zone duration. The shaded area is the total queuing delay caused by the work zone.

McCoy et al. (1980) considered user delay to be equal to the time lost while one is traveling through a construction or maintenance work zone. The time lost is taken to be a function of the difference between the average overall speed of the two-lane two-way no-passing operation and that of the normal four-lane divided operation. However, queuing delay is not taken into account in their study, because they do not consider the situation in which the approaching traffic volume exceeds the work zone capacity.

Schonfeld and Chien (1999) developed a mathematical model to optimize work zone lengths for two-lane (one lane per direction) highways where one lane in each direction at a time was closed for performing maintenance activities. In their study, deterministic queuing theory is applied to estimate user queuing delays caused
by the lane closures. In addition to the queuing delay, the moving delay incurred by vehicles traversing the work zone is also included in the user delay function.

\[ Q: \text{Demand Flow rate (vph)} \]
\[ C: \text{Roadway Capacity (vph)} \]
\[ C_w: \text{Work Zone Capacity (vph)} \]
\[ t_1: \text{Lane Closure Duration (h)} \]
\[ t_2: \text{Duration for Discharging the Queue (h)} \]
\[ \text{Queuing Delay (veh-hr)} \]

**Figure 2.1 Queuing delay estimated by the deterministic queuing model**

Most of other models have used average daily traffic (ADT) as the input data for vehicle volume. However, the daily peaking pattern can have a significant impact on average speeds and delays during the day. Memmott and Dudek (1984) developed a computer program, called Queue and User Cost Evaluation of Work Zones (QUEWZ) which employed the deterministic queuing model to estimate queuing
delay using average hourly traffic (AHT). In addition, the approaching speed was calculated using equations taken from the Highway Economic Evaluation Model (1976) and hypothetical speed-volume relations were used to estimate the moving delay through the lane-closure section. However, QUEWZ was developed based on traffic data collected from Texas highways, and is therefore inappropriate to apply to highways in other states.

Cassidy et al. (1994) presented a procedure for predicting vehicle delays at two-lane highway work zones. The procedure estimates the delay by predicting the average directional right-of-way times per cycle, occurring over a given time interval. Once these effective green times have been estimated, deterministic queuing analysis techniques are used to compute the delays.

Jiang (1999) conducted a delay study for the Indiana Department of Transportation, in which the deceleration and acceleration delays to users were also considered. Work zone related delays were classified into four categories: (1) deceleration delay, experienced due to the vehicle deceleration before entering the work zone; (2) moving delay, experienced by vehicles passing through work zones at a lower than normal traveling speed; (3) acceleration delay, experienced as the vehicle accelerates after the work zones; and (4) queuing delay, caused by vehicles arriving at a higher rate than the discharge rate. According to this study, vehicle queues may still form even when traffic flow is below the work zone capacity, because of the randomness of traffic flow.

Jiang (2001) developed a queue estimation method to calculate traffic delay
using queue-discharge rates instead of work zone capacity, having noticed that the
former are lower than the latter (Jiang, 1999).

It should be noted that traffic delay at work zones estimated using the
deterministic queuing model tends to be slightly underestimated because shock-wave
delays are neglected. Nevertheless, the deterministic queuing model is still widely
accepted by traffic engineers because of its simplicity.

2.2.2 Macroscopic Simulation Models

In addition to the macroscopic analytical methods, there are several
macroscopic software packages available for estimating queue lengths and delays at
work zones. Chitturi and Benekohal (2004) compared the performance of QUEWZ,
FRESIM and QuickZone in estimating traffic delay at work zones using field data
collected from eleven freeway work zones in Illinois. Their findings are shown as
follows:

- The results generated by QUEWZ did not match the field data. It was found that
  QUEWZ could overestimate the average speed of vehicles in both queuing and
  non-queuing sites. The queue lengths estimated by QUEWZ were far below the
  observed queue lengths. The main reason for this is the outdated speed-flow
  relationship used in QUEWZ.
- FRESIM consistently overestimated speeds under queuing conditions,
  overestimated queue lengths for half of the cases, and underestimated queue
lengths for the rest.

- QuickZone did not allow for the delays due to the slower speeds experienced in work zones. In addition, QuickZone consistently under-predicted queue lengths and delays in comparison to the field data.

### 2.2.3 Microscopic Simulation Tools

Maze and Kamyab (1999) used ARENA, a simulation model with an advanced animation module, to develop a work zone simulation model, including simplified car-following and lane changing algorithms to estimate work zone delays. Traffic delay at work zones was based on the average travel times simulated by the computer program which took into account the interactions between vehicles. However, they simply used the deterministic queuing method to estimate the delay at the upstream of merge area so that it output an underestimated delay value.

In addition to ARENA, a number of microscopic traffic simulation tools including INTEGRATION, CORSIM, PARAMCS, and VISSIM have been used by researchers to estimate negative traffic impacts at work zones. Wu (2000) applied INTEGRATION to estimate work zone traffic delays. To determine the INTEGRATION’s performance, a comparison of its accuracy against that of FRESIM was made. The results showed that the traffic delays estimated by INTEGRATION gave a better fit to the field data. Unfortunately, INTEGRATION is only applicable when the approaching traffic volume is less than the work zone capacity.
Chien et al. (2002) proposed a method for calculating traffic delay by integrating simulation data obtained from CORSIM with the deterministic queuing model. In their study, this simulation-based model was applied to estimate the average queuing delays at different traffic congestion levels. However, their model requires extensive calibration and validation of CORSIM.

Chan (2002) used the microscopic traffic simulation software, PARAMICS to estimate the traffic delay due to lane closures during maintenance activities. However, this model is again only applicable when there is no queue, that is, the approaching traffic volume is less than the work zone capacity.

Yang et al. (2008) developed a hybrid method that integrates macroscopic analytical methods with microscopic simulation to calculate work zone delays. In unsaturated traffic conditions, the microscopic simulation tool, CORSIM, is used to estimate delays because a microscopic simulation process is supposed to estimate delays more precisely than macroscopic analytical methods. However, traffic delay at work zones could be underestimated in saturated and oversaturated conditions, that is, CORSIM may be unable to estimate delays precisely in conditions that cause queue spillbacks to the vehicle entry nodes. Hence, a classical deterministic queuing model is employed for these conditions.

Although these microscopic traffic simulation tools are able to model complex traffic dynamics at the individual vehicle level and provide precise results, greater time and effort is required from the users. In addition, some traffic simulation software (e.g., CORSIM) cannot accurately represent driver behavior when drivers are
approaching to a work zone, because they simulate a work zone as a prolonged incident blockage. When modeling a lane blockage in CORSIM, the software assumes that drivers have no knowledge of the approaching blockage and that there is no transition taper. Although some other software (e.g., INTEGRATION) do a better job of capturing appropriate lane changing behavior at work zones, they do not allow users to modify the work zone configuration. However, work zone configuration and driver behavior must be taken into account, because these factors have a significant impact on traffic delay at work zones.

2.2.4 Cellular Automata Models

As mentioned above, the existing microscopic simulation tools, such as CORSIM and VISSIM, are not capable of accurately addressing the dynamics that take place within work zones, because their car-following models and lane changing algorithms are based on normal or near normal traffic flow conditions. These tools do not allow users to incorporate any external logic, which is necessary to simulate the impact of the work zone configuration and driver acceleration-deceleration behavior on traffic delay at work zones. They also require high levels of computational resources and time (Edara and Cottrell, 2007).

Cellular Automata (CA) models have been used extensively for freeway traffic simulation and also for pedestrian traffic simulation (Blue and Adler, 2001), railway traffic simulation (Li et al., 2005) and intersection traffic simulation (Spyropoulou,
2007). Cremer and Ludwig (1986) first proposed a CA model for vehicular traffic. Nagel and Schreckenberg (1992) developed a traffic flow model based on the CA concept, which was found to have computational advantages in modeling complex systems. To capture important traffic features, Schadschneider and Schreckenberg (1997) made some modifications to Nagel and Schreckenberg’s CA model. Brake light models have also been developed on the basis of Nagel and Schreckenberg’s CA model, reflecting the fact that lag vehicle behavior is dependent on the brake light status of its leading vehicle (Knospe et al., 2000; Kerner et al., 2002). Bham and Benekohal (2004) proposed a CELLSIM model based on CA and car-following concepts to simulate the single-lane highway traffic. Larraga et al. (2005) presented a modified CA model, altering the deceleration rule to simulate traffic in a single-lane highway with a ring topology.

Since the above CA models were all developed for single-lane traffic, which cannot represent real-world multi-lane traffic, various multi-lane CA models have also been developed (Nagatani, 1994; Rickert et al., 1996; Wanger et al., 1997; Jia et al., 2005; Nassab et al., 2006). For instance, Rickert et al. (1996) proposed a simple two-lane CA model.

However, the above-mentioned multi-lane CA models are based on the homogeneous traffic flow, while real-world traffic is usually heterogeneous. One method of solving this problem is to convert different vehicle types into a standard vehicle type using an equivalency factor, allowing the homogeneous CA models to be applied. Nevertheless, the main drawback of this method is that it is difficult to
interpret the interactions between different types of vehicles. A heterogeneous CA model would be more efficient. Some other researchers (Chowdhury et al., 1997; Lan and Chang, 2005; Mallikarjuna and Rao, 2009) have attempted to develop such models. Chowdhury et al. (1997) presented a two-lane CA model with two different types of vehicles, characterized by two different values for their maximum speed. Lan and Chang (2005) modeled heterogeneous traffic comprising cars and two-wheeled vehicles with their improved CA model. However, their model was deterministic and no randomization rule was applied to reflect stochastic driver behavior. Mallikarjuna and Rao (2009) modified the basic structure of the CA model in order to incorporate heterogeneous traffic flow characteristics.

In multi-lane CA models, it is also necessary to consider vehicle lane changing behavior. Rickert et al. (1996) and Wanger et al. (1997) introduced a set of lane changing rules to characterize lane changing behavior in a two-lane road without lane reductions. To simulate work zone traffic, Nassab et al. (2006) applied symmetric lane changing rules to model vehicles’ lateral movements near a partial lane closure. However, the length of the transition area was assumed to be two cells in their study, which was inconsistent with real-world work zone configurations. Another deficiency in their model is that a lane change maneuver is assumed to be completed in one second. This is inconsistent with the field observation that it takes at least two seconds for a vehicle to complete a lateral movement in work zone. Overall, therefore, more realistic lane changing rules for heterogeneous work zone traffic should be introduced.
2.3 Work Zone Operational Strategy

Given a work zone project, contractors may choose either a single work zone operational strategy or a subwork zone operational strategy (Martinelli and Xu, 1996). A single work zone operational strategy closes the entire lane in one travel direction during the project implementation. In this strategy, the activity area length of the work zone is the total length to be maintained or constructed. A subwork zone operational strategy divides the entire work zone into multiple subwork zone segments, and the relevant work activities are undertaken in one segment at a time.

2.3.1 Single Work Zone Operational Strategy

Over the past several decades, various models have been developed to determine an optimal single work zone operational strategy (Schonfeld and Chien, 1999; Chen and Schonfeld, 2006; and Yang et al., 2008).

Some of these models aim to identify an optimal single work zone operational strategy without taking into account the project start time. For example, Schonfeld and Chien (1999) established a one-dimensional unconstrained minimization model that minimized the average work zone cost by determining the optimum single work zone length for two-lane highways. This interesting work was later extended by Chien and Schonfeld (2001) to a four-lane highway work zone with one lane closure. Chen et al. (2005) and Chen and Schonfeld (2006) continued exploring the optimum single work zone length issue for two-lane and four-lane highways with alternative routes.
Tien and Schonfeld (2006) examined the same issue by incorporating tradeoffs between work duration and cost.

However, project start time must be considered in a work zone project because different project start times may result in different total work zone costs, owing to the variations in traffic flow. Jiang and Adeli (2003) therefore took into account the start time choice of a work zone project and built a constrained minimization model to simultaneously identify the optimum single work zone length and the best project start time. Their objective was to minimize the average work zone cost. It should be pointed out that their user delay cost estimation formula has to be corrected when the queue vanishes within a time interval. Yang et al. (2008) further enhanced the model and improved the solution algorithm by estimating total traffic delay using a hybrid deterministic queuing model and microscopic simulation, which is similar to the traffic delay estimation method proposed by Chien et al. (2002).

However, it should be noted that the models for the single work zone operational strategy are based on the assumption that the length of the entire work zone is unknown. Obviously, this assumption is not consistent with real-world work zone projects. In reality, the location and the length of the entire work zone are always given to contractors. Thus, these models cannot be applied to real work zone projects.

### 2.3.2 Subwork Zone Operational Strategy

In contrast to single work zone operational strategies, subwork zone
operational strategies can be applied to the real-world work zone project with a known total length. Up till now, only a limited number of research studies are found in relation to the subwork zone operational strategy.

McCoy et al. (1980) contributed an analytical solution for the optimum subwork zone length, resulting in the minimal total work zone cost for two-lane two-way no-passing traffic operations in work zone projects on rural four-lane divided highways. However, they overlooked the choice of project start time and used the average daily traffic (ADT) volumes, which do not capture the real variations in traffic. Their research work was thus improved by Martinelli and Xu (1996) and McCoy and Mennenga (1998). In these two studies, user delay was estimated based on the average hourly traffic (AHT) volumes. Unfortunately, they obtained the optimal solution by setting the derivative of the total work zone cost function equal to zero. This causes an incorrect result because their model is non-differentiable. More importantly, they did not take into account the time window constraints for the subwork zone strategy. To deal with this problem, Tang and Chien (2008) adopted the simulated annealing algorithm and genetic algorithm to find the optimal subwork zone strategy subject to time window constraints.

2.4 Vehicle Crash Risk Assessment at Work Zones

A number of studies have been conducted to assess work zone risk (Hall and Lorenz, 1989; Ha and Nemeth, 1995; Zhao 2001). They conclude that work zones
produce a significantly higher rate of crashes than non-work zone locations. In particular, Hall and Lorenz (1989) found a 26% increase in motor vehicle crashes during road construction or maintenance. Garber and Woo (1990) found that, on average, accident rates increased by approximately 57% in work zones on two-lane urban highways. Zhao (2001) investigated the characteristics of work zone crashes in Virginia for the years 1996-1999 and pointed out that a higher proportion of crashes are fatal in work zones than in other locations. These facts underscore the urgent need to develop a model to accurately assess work zone crash risk and examine its contributing factors.

To date, two types of approaches have been used most frequently to analyze vehicle crash risk. One is the collective approach, characterized by crash frequency modeling (Abdel-Aty and Keller, 2005). The other is the individual approach, characterized by each individual crash case. The second approach focuses on investigating the relationship between the occurrence probability of a crash of specified severity and the contributing factors of driver, vehicle, roadway, and environmental characteristics.

### 2.4.1 Collective Approach: Vehicle Crash Frequency Modeling at Work Zones

Khattak et al. (2002) applied negative binomial techniques to model work zone crash frequencies using a unique dataset of California freeway work zones that
included crash data, road inventory data (ADT and urban/rural), and work zone related data (duration, length and location). The model results showed that frequencies increased with work zone duration, length and ADT.

Qi et al. (2005) conducted a detailed investigation of rear-end crashes at work zones that occurred in New York State between 1994 and 2001. Since data were available only for work zones at which accidents had occurred, truncated count data models, including truncated Poisson and negative binomial models, were developed to investigate the relationship between the crash frequency and work zone characteristics. Eight categories of independent variables including work zone type, control device, layout, lane blockage, operation, location, facility type and traffic volume were used for the model formulation. The truncated count data model results indicated that work zones with flaggers, alternating one-way traffic and higher ADT cause more rear-end accidents.

Srinivasan et al. (2007) modeled the location of rear-end crashes within work zones as functions of the lengths of different work zone segments, traffic volume, weather and other exogenous factors. The empirical results indicated that weather conditions and traffic characteristics are statistically significant and intuitively reasonable predictors of the locations of crashes within work zones. The model was also applied to assess the relative safety of different work-zone segments in terms of the probability of a crash per lane-mile of the segment. The results indicated that the advance warning area is particularly unsafe during times of peak traffic flow and bad weather conditions (relative to off-peak and good weather). Further, the exit area was
also found to be relatively unsafe, especially during peak periods.

Kim et al. (2007) developed a modified negative binomial regression model to estimate rear-end crash risk, using accident data from Washington State. Compared with previous rear-end crash risk models, their model had two major advantages: 1) it directly considered a driver’s response time distribution; and 2) it applied a new dual-impact structure, accounting for both the probability of a vehicle becoming an obstacle ($P_o$) and the probability of the following vehicle’s reaction failure ($P_f$). The results showed that urban area, curvature, off-ramp and merge, shoulder width, and merge section are the factors influencing the rear-end crash probabilities.

Note that the above studies are based on the ability of the model formulation to capture the underlying distribution of the crash count data. Recently, some researchers (Abdel-Aty and Keller, 2005; Chang and Chen, 2005) have proposed distribution-free methodologies that are essentially driven by the observed data. These methodologies do not require inherent assumptions about the distribution of the crash data.

2.4.2 Individual Approach: Occurrence Probability of Vehicle Crashes at Work Zones

When analyzing the occurrence probability of a work zone crash and its contributing factors, most of the previous safety-related studies have used logistic regression techniques.

Dissanayake and Lu (2002) developed a set of sequential binary logistic
regression models to analyze the contributing factors and predict the crash severities of single-vehicle fixed object work zone crashes involving young drivers. The researchers utilized the Statistical Analysis Software (SAS, 2008) to develop the models and take into account crash factors such as gender, driver impairment, and the geometric conditions of crash locations.

Using data from 376 fatal work zone crashes in Texas between 1997 and 1999, Hill (2003) investigated the effectiveness of traffic counter measures, such as the use of an officer/flagman or a stop/go signal, on fatal crash risk using the logistic regression approach. The results of this study indicate that there is a significant difference in crash types and driver error between daytime crashes and nighttime crashes. Differences also exist between driver genders. In addition, commercial truck related crashes are more likely to involve multiple vehicles. Furthermore, the use of an officer/flagman or a stop/go signal was found to reduce the chance of a crash by 68% or 64%, respectively.

Harb et al. (2008) developed a conditional logistic regression model based on the Florida Crash Records Database for the years 2002-2004 to estimate work zone rear-end crash risk. According to their results, roadway geometry, weather conditions, age, gender, lighting conditions, residence code, and driving under the influence of alcohol and/or drugs are the most significant risk factors associated with work zone crashes.

Based on Kansas highway work zone crash data for the years 2003-2004, Li and Bai (2008) developed crash-severity-index models using the logistic regression technique. Their models only incorporated significant risk factors that had an impact on crash severity. Finally, they validated their models using recent crash data and analyzed their ability to assess work zone risk levels. Later, Li and Bai (2009) used
binary logistic regression to evaluate the effectiveness of the temporary traffic control (TTC) methods commonly used in work zones.

Wang et al. (2008) employed a two-stage analysis procedure to analyze fatal work zone crash characteristics, based on 421 sets of fatal work zone crashes in Florida over a four-year period (from 2002 to 2005). In the first stage, a descriptive statistical method was used to examine the distributions of fatal work zone crashes across various variables. In the second stage, a binary logistic regression was adopted to develop models to predict the occurrence probabilities of fatal crashes at work zones by crash type, age group, and predominant contributing factors. Their results showed that the probability of a fatal work zone crash for middle-aged drivers (25-64 year-old) increases when heavy vehicles and/or alcohol are involved. For elderly drivers, the presence of an intersection, bridge, ramp or road access is a significant factor increasing the probability of a fatal work zone crash.

In addition to the conventional logistic regression techniques, more advanced logit-based approaches, such as the nested logit model (Shankar et al., 1996) and the mixed logit model (Milton et al., 2008) have also been employed to analyze the same issue. For example, Abdel-Aty and Abdelwahab (2004) proposed a nested logit model to estimate the probability of a car-truck rear-end crash. To develop an appropriate nesting structure, many possible structures were considered and calibrated. Using the likelihood ratio index and classification accuracy of the test sets as measures of goodness-of-fit, a two-level nesting structure, with car-truck in one group and all other combinations (car-car, truck-car, and truck-truck) in another group, was finally
selected using the general estimates system (GES) databases. The model showed the
significant variables to be driver’s age, gender, traffic control device, action initiated
by the lead vehicle and the inattention and vision obstruction of the driver in the
following vehicle.

2.5 Identified Research Limitations and Gaps

It can be seen from the literature review that there are several limitations and
gaps in the existing studies, implying the need for further research on work zone
traffic delay estimation, operational strategy analysis and risk assessment. This section
highlights these limitations and gaps.

2.5.1 Limitations in Traffic Delay Estimation at Work Zones

The existing literature shows that the microscopic traffic simulation methods
provide more accurate traffic delay estimates than the macroscopic methods. Among
the microscopic simulation methods, the CA model is a promising and effective model
for traffic delay estimation because it has the ability to eliminate the limitations that
exist in the current microscopic simulation tools. However, the existing CA models
cannot be applied directly to estimate work zone traffic delays for the following two
reasons.

First, there are two flaws in the previous CA models. The randomization
probability parameter in the CA model reflects the likelihood of a driver increasing or decreasing the traveling speed. However, this important parameter was set to a hypothetical constant value in the previous CA models, resulting in a “coarse” description of traffic operations. In reality, the randomization probability should vary with traffic flow and work zone configurations (Al-Kaisy and Hall, 2001; Lee et al., 2004). The other flaw is that the previous CA models all assume that a vehicle can complete a lane change in one second. However, such an assumption is inconsistent with reality as this usually takes a vehicle at least two seconds (Gundaliya et al., 2008).

Second, the rules adopted by the conventional CA models in previous studies are inadequate to describe the complex dynamics at work zones, because of the unique work zone traffic characteristics. Therefore, supplementary rules should be added in order to replicate realistic work zone driver behavior.

2.5.2 Gaps in Work Zone Operational Strategy Studies

From the literature, it can be seen that a subwork zone strategy is more realistic than a single work zone strategy. Although a number of models have been proposed to find an optimal subwork zone strategy, they require improvements or reformulation in view of the following three issues. First, these models may overestimate the total queuing delay because they neglect to account for the fact that a queue may completely disappear before the end of a time interval, when the arriving
traffic flow is less than the work zone capacity for that time interval. In the estimation of moving delay, these models do not consider that the vehicle departure rate is equal to the work zone capacity before the queue completely disappears. Second, they apply an unrealistic constant approaching traffic speed. According to HCM (2000), traffic speed actually varies with traffic volume. Third, they do not take into account the uniform subwork zone length constraint. However, each subwork zone usually has a uniform length in reality. Hence, it is important and would be more practical to find a subwork zone operational strategy considering variable traffic speed, time window and uniform subwork zone length constraints.

It should also be noted that the existing subwork zone operational strategy studies all aim to minimize the total work zone cost from the systemic perspective. However, work zone contractors have no interest in minimizing the total work zone cost. In fact, they are mainly concerned with the minimizing the maintenance or construction cost. Hence, it would also be interesting and important to develop a new model which aimed to minimize the maintenance/construction cost subject to queue length and travel delay constraints from the work zone contractor’s standpoint. However, little effort has been made in this direction.

2.5.3 Limitations of Work Zone Crash Risk Assessment Studies

Although work zone crash risk assessment has been attempted from various perspectives, there are several potential problems and gaps in the existing risk
First, the existing literature suggests that the logistic regression technique and its variations are useful for identifying the crash risk factors and predicting work zone crash probabilities. However, it should be pointed out that the accuracy of the result will be highly dependent on the quality of the historical accident data. Inaccurate or biased results can sometimes occur due to poor quality of historical data, caused by traffic police recording accidents incorrectly (Kamalasudhan et al., 2002). There is also a probability that no historical accident data will be available, especially for a newly-built road or a newly-proposed work zone. In this case, the crash risk cannot be evaluated. Hence, other types of data, such as traffic data, should be used to assess work zone crash risk. Many studies (Hu et al., 2004; Hourdos et al., 2006; Oh et al., 2010) have already provided adequate evidence that modes based on traffic data can yield reliable and accurate risk estimates.

Second, almost all of the existing work zone crash risk assessment studies focus on separately assessing the occurrence frequency and the severity of work zone crashes. However, this separate assessment cannot completely reflect a facility’s risk, rendered by a broad range of accidents from frequent-minor to rare-major accidents. In practice, traffic safety engineers seem to place much more concern on the vehicle occupant’s casualty risk, e.g., the likelihood of a driver/passenger being killed or injured in a work zone, and on the relationship between the frequencies and consequences of work zone crashes. In this case, a quantitative risk assessment (QRA) model should be introduced in order to compensate for the weak points in the current
Chapter 2 Literature Review

studies—the QRA model can provide a numerical evaluation of accident consequences, frequencies, and their combination into an overall risk measure (Borysiewicz et al., 2006).

2.6 Summary

This chapter has presented a critical literature review focusing on the three major work zone operational issues: traffic delay estimation, operational strategy analysis and risk assessment for work zones. Arising from the literature, several potential problems and gaps regarding each issue have been identified.

Firstly, as pointed out in Section 2.3, the existing CA models cannot be applied to estimate work zone traffic delay due to the unique work zone traffic characteristics though they outperform the microscopic simulation methods. Therefore, there is a practical need to develop a specific CA model to better simulate work zone traffic.

Secondly, current methodologies cannot provide an accurate estimate of work zone capacity, which is a key input in the optimal subwork zone operational strategy problem. Hence, a decision tree-based model will be proposed in the next chapter to accurately estimate work zone capacity.

Thirdly, there are several limitations and flaws in the existing models for subwork zone operational strategies. It is hence necessary to propose a new model to remedy these. So far, the existing models all take a systemic perspective, minimizing the total work zone cost. Little effort has been made to minimize
maintenance/construction costs, which work zone contractors are mainly concerned.

Finally, the existing work zone crash risk assessment models rely highly on historical accident data, meaning that it is not able to assess crash risk for a newly-proposed work zone for which the historical data is unavailable. Some other methods are required for such cases. Although traffic safety engineers are concerned with vehicle occupant’s casualty risk, which is a combination of the occurrence frequency and the consequence of work zone crashes, there is a lack of studies assessing this. The QRA technique is a promising approach in dealing with this problem because it not only provides a numerical evaluation of accident consequences and frequencies, but also combines them into an overall risk measure.
CHAPTER 3  HETEROGENEOUS CELLULAR AUTOMATA MODEL FOR WORK ZONE TRAFFIC DELAY ESTIMATION

3.1 Introduction

The presence of work zones would reduce road capacity and increase traffic delay due to lane closures. The accurate estimation of traffic delay occurred in a work zone is of utmost importance for traffic engineers because the estimation results would affect the efficiency of work zone plans and traffic management strategies. Microscopic traffic simulation is an efficient method to provide an accurate estimate of traffic delay occurred in work zone. Maze and Kamyab (1999) investigated that microscopic simulation tools cannot describe the interaction between vehicles and work zone configurations well because the work zone is simulated through a prolonged incident blockage with no transition area. In addition, these traffic simulation tools (e.g., CORSIM) require high computational resources and long execution time (Bham and Benekohal, 2004).

Numerous traffic flow models, such as CELLSIM (Bham and Benekohal, 2004) and TRANSIMS, have been developed based on the CA model and have exhibited high computational efficiencies. Most of these CA-based models are only applicable to simulate homogeneous traffic flow. Obviously, they cannot describe the dynamics and interactions of vehicles of different types in work zone. In CA models, the randomization probability parameter reflects the likelihood of a driver speeding up or slowing down the traveling speed. It can be used to describe stochastic driver acceleration-deceleration behavior. However, this important parameter was set to a
hypothetical constant value in the previous CA-based models so that they can only provide a “coarse” description of traffic operations. In reality, the value of randomization probability is not a constant value for work zone traffic. It varies with traffic flow and work zone configuration factors including the activity area length and the transition area length (Lee et al., 2004; Heaslip et al., 2008). In addition, work zone forbids lane changes from the through lane to the closed lane in the transition area. Therefore, a heterogeneous cellular automata (HCA) model should be developed to simulate the heterogeneous work zone traffic.

In the HCA model, the randomization probability parameter should be modeled as a function of the activity area length, the transition area length and the volumes of different types of vehicles traveling through work zone. In addition, the HCA model will add a new lateral speed and position updating rule to account for the fact that it usually takes more than two seconds for a vehicle to complete a lane change maneuver in reality. The lateral speed and position updating rule would ensure that the simulation of vehicle’s lateral movement in work zone is close to reality. The developed HCA model can further be applied to estimate traffic delay occurred in work zone.

### 3.2 Model Development

#### 3.2.1 Work Zone Configuration

In reality, the lane width varies from 3.3 m to 3.6 m in Singapore. For simplicity, it is assumed that each lane has the uniform width of 3.5 m in this study. Figure 3.1 depicts a work zone located on a three-lane one-way road. It is assumed
that the work zone causing one partial lane reduction is of length \( L = L_a + 2L_t \). It is placed between the positions \( x_1 \) and \( x_4 \) on Lane 1, as shown in Figure 3.1. Obviously, the parameters \( L_a \) and \( L_t \) can influence the driver acceleration-deceleration behavior.

![Figure 3.1 Schematic representation of work zone in simulation](image)

### 3.2.2 Cell and Vehicle Sizes

In the previous CA models, researchers assumed that each cell has a uniform length of 7.5 m and each cell width is equal to the lane width. Each vehicle has the uniform size and occupies one cell. Obviously, such a definition of cell size in their CA models was too coarse, which led to unrealistic acceleration and deceleration rates. The uniform vehicle size cannot account for the fact that different types of vehicles have different vehicle sizes.

Although the smaller cell size can represent different types of vehicles more accurately, it requires a huge memory for computation. Therefore, a fine cell size should be defined so that it is not only able to capture different types of vehicles’ dimensions but also reduces the computational resources. In this chapter, such a fine
cell size is determined based on the vehicle dimensions. According to the actual sizes of different vehicle types listed in Table 3.1, the length and width of a fine cell are taken as 0.5 m and 0.7 m, respectively. Considering the clearance of two stopped vehicles, the vehicle size represented in this chapter should be slightly larger than the actual size. Therefore, a light vehicle (e.g., car) occupies 9×3 cells and a heavy vehicle (e.g., truck) takes away 16×4 cells. The comparison of vehicle sizes taken in this study versus actual sizes is shown in Table 3.1. It can be seen that the physical representation of the vehicle in this study is very close to the reality.

<table>
<thead>
<tr>
<th>Vehicle Type</th>
<th>Length (Longitudinal)</th>
<th></th>
<th>Width (Lateral)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual (m)</td>
<td>HCA (cell/m)</td>
<td>Clearance (m)</td>
</tr>
<tr>
<td>Light (e.g. car)</td>
<td>4.2</td>
<td>9/4.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Heavy (e.g., truck)</td>
<td>7.2</td>
<td>16/8.0</td>
<td>0.8</td>
</tr>
</tbody>
</table>

### 3.2.3 Vehicle Updating Procedure

When vehicles move over the multi-lane road space, drivers have to make a decision on the vehicles’ longitudinal and lateral movements in the next time step. Here, a time step is equal to one second in the HCA model. Therefore, the vehicle updating procedure implemented by the HCA consists of two parts. The first part updates vehicles’ longitudinal movements based on the forwarding rules and the second part updates vehicles’ lateral movements using the lane changing rules. Vehicles can be generated either by the negative exponential distribution or by the distribution of the observed arrival pattern. The generated vehicles will travel from the
leftmost to the rightmost of road.

### 3.2.3.1 Updating vehicles’ longitudinal movements

In free traffic conditions, there are few vehicle interactions and drivers independently make their decisions. However, when the time headway to the leading vehicle is shorter than a critical value (in the congestion conditions), the following vehicle’s longitudinal movement will be influenced by the leading vehicle behavior.

In this context, the brake light (BL) model proposed by Knospe et al. (2000) deserves much attention because it uses interaction time headway and the status of brake light to determine whether there are vehicle interactions. Therefore, the forwarding rules in the HCA model are similar to the previous BL rules but with the following three modifications due to the unique traffic characteristics in work zone:

1. Driver acceleration-deceleration behavior varies with the traffic flow and work zone configuration. In this study, the randomization probability outside work zone \((p_1)\) is dependent on the light vehicle flow \((f_L)\) and heavy vehicle flow \((f_H)\). The randomization probability in work zone \((p_2)\) should be modeled as a function of the factors of the light vehicle flow \((f_L)\), the heavy vehicle flow \((f_H)\), the activity area length \((L_a)\) and the transition area length \((L_t)\).

2. In reality, traffic engineers usually post speed limits at the advance warning sign (position \(x_0\)) to limit the vehicle’s traveling speed. Therefore, the allowed maximum speed for vehicles traveling in work zone should not be larger than the posted speed limit \(V_{pos}\).
(3). The available front gap for the front vehicle in the blocked lane depends on the vehicle position and work zone configuration.

### 3.2.3.1.1 Notation

- \((x_n^t, y_n^t)\): the top left position of vehicle \(n\) at time \(t\);
- \(v_n^t\): the speed of vehicle \(n\) at time \(t\), where \(v_n^t = 0, 1, 2, \ldots, v_{max}^k\);
- \(V_{pos}\): the posted speed limit;
- \(v_{max}^k\): the maximal speed of vehicle type \(k\) (1 for light vehicle; 2 for heavy vehicle);
- \(v_{max}^n\): the maximal speed of vehicle \(n\);
- \(b_n^t\): the brake light status of vehicle \(n\) at time \(t\) (0 if the brake light is off; 1 if the brake light is on);
- \(l_n(k)\): the vehicle length of vehicle \(n\);
- \(d_n^t\): the available front gap between the vehicle \(n\) and its leading vehicle, where \(d_n^t = x_{n-1}^t - x_n^t - l_{n-1}(k)\). Note that the available front gap of vehicle \(n\) in Lane 1 would be \(d_n^t = x_i - x_n^t - y_n^t \cdot L / 5\) as it approaches the transition area.
- \(t_n^h\): the available time headway of vehicle \(n\) at time \(t\), where \(t_n^h = d_n^t / v_n^t\);
- \(a_n(v_{n-1}^t, k)\): the acceleration rate of vehicle \(n\) of vehicle type \(k\) at time \(t\);
- \(d_n(k)\): the deceleration rate of vehicle \(n\) of vehicle type \(k\);
- \(d_{security}\): the security longitudinal distance between two vehicles;
- \(p_1\): the randomization probability outside work zone;
$p_2$: the randomization probability in work zone;

$f_L$: the light vehicle flow;

$f_H$: the heavy vehicle flow;

$L_a$: the activity area length of work zone;

$L_t$: the transition area length of work zone.

### 3.2.3.1.2 Forwarding rules

The following consecutive forwarding rules are performed in parallel for all vehicles to update vehicles’ longitudinal speeds and positions:

**Rule 0:** *(Determination of randomization probability and maximum speed).*

Outside work zone (e.g., $x_n' < x_0$ or $x_n' > x_4$),

$$ p = p_1(f_L, f_H), \quad v_n^{\text{max}} = v_k^{\text{max}}; $$

In work zone (e.g., $x_0 \leq x_n' \leq x_4$),

$$ p = p_2(f_L, f_H, L_a, L_t), \quad v_n^{\text{max}} = \min\{V_{\text{pos}}, v_k^{\text{max}}\}. $$

**Rule 1:** *(Acceleration).* If the available time headway $t_n^{b}$ is greater than the interaction headway $t_s$ or the status of brake lights for the vehicle $n$ and its leading vehicle both are equal to zero, namely, if $(b_{n-1}^{i} = 0$ and $b_{n-1}^{i-1} = 0)$ or $(t_n^{b} > t_s)$, then

$$ v_n' = \min\{v_n^{i-1} + a_n(v_n^{i-1}, k), v_n^{\text{max}}\}. $$

**Rule 2:** *(Deceleration).* Calculate the effective available front gap after considering the anticipated speed of the leading vehicle:

$$ d'_{n,\text{eff}} = d_n' + \max\left\{ \min\left\{v_n^{i-1}, d_n'^{-}\right\} - d_{\text{security}}, 0 \right\} $$

(3.1)

If $d'_{n,\text{eff}} < v_n'$, the speed of vehicle $n$ is reduced to $d'_{n,\text{eff}}': \quad v_n' = \min(v_n', d'_{n,\text{eff}}')$. If $v_n' < v_n'^{i-1}$, the brake light status of vehicle $n$ is updated as 1, namely, $b_n' = 1$. 

52
Rule 3: (Randomization). Update the speed of vehicle \( n \): \[ v'_n = \max(v'_n - d'_n(k), 0) \]
with the randomization probability \( p \).

Rule 4: (Vehicle movement). Each vehicle moves forward according to its new speed,
\[ x'_n = x''_n + v'_n. \]

Note that all vehicles update their speeds and positions simultaneously. In addition, the right most vehicles update first and then the next one will be taken into consideration.

3.2.3.2 Updating vehicles’ lateral movements

The lane changing rules for the HCA model employ two criteria including incentive criterion and safety criterion that are used in the previous CA models. According to the unique work zone traffic characteristics, the HCA model also adds three supplementary rules including one lateral speed and position updating rule and two lane change constraint rules in work zone.

(a) Incentive criterion
1. \( v'_n > d'_n \) and \( v'_n > v'_{n-1} \);
2. \( d'_n < d'_{n,n,f} \). Here \( d'_{n,n,f} \) is the available front gap between the vehicle \( n \) and its front neighboring vehicle in the target lane.

(b) Safety criterion
3. \( d'_{n,b,b} > v'^{n,n,b}_{\text{max}} \), where \( d'_{n,b,b} \) is the available back gap between the vehicle \( n \) and its back neighboring vehicle in the target lane, \( v'^{n,n,b}_{\text{max}} \) is the maximum speed of
its back neighboring vehicle. This criterion ensures that the vehicle would not collide with the lag vehicle in the target lane.

If the vehicle perceives that the above criteria are both satisfied, then it will make a lateral movement decision with a specified probability.

(c) Lateral speed and position updating rule

4. In reality, it usually takes a vehicle at least two seconds to complete a lateral movement. According to Chovan et al. (1994), the maximal lateral speed is not larger than 1.0 m/s (2 cell/sec). Therefore, a realistic lateral speed and position updating rule is adopted so that the vehicle’s lateral movement is close to the reality, shown as follows:

\[
y'_n = \begin{cases} 
\min \{y'^{-1}_n + 2,5 \times TL_n - l_n(k)\} & \text{if changing to the right} \\
\max \{y'^{-1}_n - 2,5 \times TL_n - l_n(k)\} & \text{if changing to the left} \\
5 \times TL_n - l_n(k), & \text{if no change}
\end{cases}
\] (3.2)

where \( TL_n \) is the lane number of the target lane that vehicle \( n \) expects to move into.

(d) Lane change constraint rules in work zone

5. The above lane changing rules can precisely describe lane changing behavior in the activity area. However, an additional rule should be added in the advance warning area because, in a simultaneous update, it is possible in this area that a vehicle from Lane 1 and a vehicle from the Lane 3 go to the overlapped cells in Lane 2, resulting in a lateral collision. In order to avoid this collision, the HCA model chooses a vehicle at random and then allows it to perform its requested lane change in the advance warning area.
6. Since vehicles in Lane 1 are ultimately redirected to Lane 2, they are forbidden to take lane changes from Lane 2 to Lane 1 in the transition area. Similarly, it is not allowed for these vehicles to take lane changes into Lane 2 in the termination area.

### 3.3 Model Calibration

The HCA model requires inputs to its parameters to simulate the heterogeneous work zone traffic. The maximum speed, acceleration rate, deceleration rate and the randomization probability are the four important model parameters. In order to replicate the realistic heterogeneous work zone traffic, the HCA model must be calibrated. 41 sets of expressway work zone data and 15 sets of arterial work zone data, respectively collected on the six-lane two-way PIE expressway and arterial roads in Singapore from August 1997 to March 1998, are used to calibrate these four input parameters. Among the collected datasets, 21 sets from expressway work zone and 9 sets from arterial work zone data are used to calibrate the HCA model. The remaining datasets are used for model validation.

Each work zone dataset consists of light vehicle flow rate, heavy vehicle flow rate, activity length and transition length, traffic speeds in and outside work zone, work zone travel time and traffic delay. Traffic speeds in and outside of work zone are captured by two laser guns at activity and advance warning areas, respectively. Two license plates are respectively located at the start of advance warning area and at the
end of termination area. Work zone travel time can be obtained by taking the
difference between these two license plate timings. Traffic delay incurred in work
zone is defined as the difference between travel times in and outside work zone.

### 3.3.1 Maximum Speed, Acceleration and Deceleration Rates

In the PIE expressway, the observed maximum travelling speed is about 109
km/h (≈60 cell/sec) outside of the work zone and 80 km/h ($V_{pos}$≈45 cell/sec) in work
zone. Similarly, the maximum speeds outside and in arterial work zones are
respectively set to 45 cell/sec and 35 cell/sec for the HCA model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>Light vehicle ($k=1$)</th>
<th>Heavy vehicle ($k=2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceleration rate*</td>
<td>$v_{n}^{t-1}$ ≤ 11 cell/sec</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>$a_{n}(v_{n}^{t-1}, k)$ (cell/sec²)</td>
<td>11 cell/sec &lt; $v_{n}^{t-1}$ ≤ 22 cell/sec</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$v_{n}^{t-1}$ &gt; 22 cell/sec</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Deceleration rate*</td>
<td>$d_{n}(k)$ (cell/sec²)</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Interaction headway* $t_s$ (sec)</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Security distance# $d_{securiy}$ (cell)</td>
<td>9</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Maximum speed $v_{max}^k$ (cell/sec)</td>
<td>Expressway</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Expressway work zone</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Arterial road</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Arterial work zone</td>
<td>35</td>
<td>35</td>
</tr>
</tbody>
</table>

$v_{n}^{t-1}$ is the traveling speed of vehicle $n$ at the start time of $t$;

*source from Mallikarjuna and Rao (2007);
#source from Mallikarjuna and Rao (2009).

Mallikarjuna and Rao (2007) investigated that the acceleration and
deceleration rates both vary with vehicle types. In addition, the vehicle’s acceleration rate is also affected by the vehicle’s traveling speed (Bham and Benekohal, 2004). For the HCA model, the acceleration and deceleration rates provided by Mallikarjuna and Rao (2007, 2009) are used in this chapter, shown in Table 3.2.

### 3.3.2 Randomization Probability

The randomization probability can be calibrated by minimizing the square error between the estimated average travel time and the observed average travel time using a trial-and-error method. The procedures are shown as follows:

1. **Step 0. (Initialize)** let \( p^{(k)} = p_0 \), where \( 0 \leq p_0 \leq 1 \), \( k = 0 \);

2. **Step 1. (Obtain the estimated travel time)** input \( p^{(k)} \) to the HCA model, and obtain the corresponding estimated average travel time \( T_s^{(k)} \);

3. **Step 2. (Update the randomization probability):** compare \( T_s^{(k)} \) and \( T_f \), where \( T_f \) is the observed average travel time, if \( T_s^{(k)} = T_f \), go to Step 3; if \( T_s^{(k)} > T_f \), let

   \[ p^{(k+1)} = p^{(k)} - \Delta p \]

   where \( \Delta p \in [0,1] \);

   if \( T_s^{(k)} < T_f \), let

   \[ p^{(k+1)} = p^{(k)} + \Delta p \]

   where \( \Delta p \in [0,1] \);

4. **Step 3. (Verify a stopping criterion):** let \( \xi \) represents the accepted error, if

   \[ (T_s^k - T_f)^2 < \xi^2 \]

   stop and output \( p^{(k)} \); otherwise, set \( k = k+1 \) and go to Step 1.

The proposed HCA model is stochastic so that the estimate from the simulation is not always the same for a particular randomization probability. To minimize the stochastic errors and obtain relative stable results, it is necessary to perform simulation more than once. Nevertheless, the more simulation runs, the simulation time required is longer. Therefore, it is preferred to find the particular simulation runs, which not only meets the precision of results but also does not
increase the simulation time greatly. Based on the theory of probability and statistics, the following equation can be used to estimate the required number of runs:

\[ N_{\text{run}} = \left( \frac{z_{\alpha/2} \sigma}{E} \right)^2 \]  

(3.3)

where \( N_{\text{run}} \) represents the required number of simulation runs, \( \sigma \) is the sample standard deviation, \( z_{\alpha/2} \) is the threshold value for \( 100(1-\alpha) \) percentile confidence interval and \( E \) represents the allowed error range.

Using the initial \( N_{\text{run}}=20 \), the maximum value of \( \sigma / E \) can be obtained among all work zones. The required number of simulation runs can be obtained after substituting the maximum \( \sigma / E \) into Eq. (3.3). Assuming that the accepted error range is 2.5% and \( \alpha = 95\% \) in expressway work zones, the value of \( \sigma / E \) is maximal as the activity area length \( L_a = 850 \) m, the light vehicle flow \( f_L = 2072 \) vph, the heavy vehicle flow \( f_H = 806 \) vph and the corresponding \( p \) is 0.031. Using Eq.(3.3), the required number of simulation runs in expressway work zones is:

\[ N_{\text{run}} = \left( \frac{1.96 \times 4.123}{61.2 \times 2.5\%} \right)^2 \approx 28 \]

In arterial work zones, it is assumed that the accepted error range is also 2.5% and \( \alpha = 95\% \). As the activity area length is 170 m, the transition area length \( L_t = 20 \) m the light vehicle flow \( f_L = 1855 \) vph and the heavy vehicle flow \( f_H = 313 \) vph, the corresponding \( \sigma / E \) is the maximum. The required number of simulation runs in arterial work zones is:

\[ N_{\text{run}} = \left( \frac{1.96 \times 0.978}{14.44 \times 2.5\%} \right)^2 \approx 29 \]

The results of the calibrated randomization probability outside and in work zones are shown in the columns of “\( p_1 \)” and “\( p_2 \)” of Table 3.3-3.4. It can be seen that
the randomization probability in expressway work zones varies from 0.028 to 0.126.

In arterial work zones, the randomization probability is relatively high, ranging from 0.176 to 0.243.

### Table 3.3 Calibrated randomization probabilities in expressway work zones

<table>
<thead>
<tr>
<th>#</th>
<th>$L_a$ (m)</th>
<th>$f_L$ (vph)</th>
<th>$f_H$ (vph)</th>
<th>$T_{wz}$ (sec/veh/zone)</th>
<th>$T_0$ (sec/veh/zone)</th>
<th>$D_{wz}$ (sec/veh/zone)</th>
<th>$p_1$</th>
<th>$p_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>330</td>
<td>1889</td>
<td>698</td>
<td>34.7</td>
<td>31.1</td>
<td>3.6</td>
<td>0.320</td>
<td>0.090</td>
</tr>
<tr>
<td>2</td>
<td>330</td>
<td>1804</td>
<td>702</td>
<td>34.1</td>
<td>30.7</td>
<td>3.4</td>
<td>0.324</td>
<td>0.113</td>
</tr>
<tr>
<td>3</td>
<td>330</td>
<td>1546</td>
<td>762</td>
<td>33.5</td>
<td>30.5</td>
<td>3</td>
<td>0.335</td>
<td>0.120</td>
</tr>
<tr>
<td>4</td>
<td>380</td>
<td>1509</td>
<td>849</td>
<td>35.5</td>
<td>32.3</td>
<td>3.2</td>
<td>0.325</td>
<td>0.116</td>
</tr>
<tr>
<td>5</td>
<td>380</td>
<td>1643</td>
<td>671</td>
<td>36.5</td>
<td>32.8</td>
<td>3.7</td>
<td>0.351</td>
<td>0.126</td>
</tr>
<tr>
<td>6</td>
<td>510</td>
<td>1943</td>
<td>793</td>
<td>44.8</td>
<td>39.4</td>
<td>5.4</td>
<td>0.305</td>
<td>0.070</td>
</tr>
<tr>
<td>7</td>
<td>510</td>
<td>1938</td>
<td>830</td>
<td>45.8</td>
<td>39.7</td>
<td>6.1</td>
<td>0.304</td>
<td>0.068</td>
</tr>
<tr>
<td>8</td>
<td>510</td>
<td>1968</td>
<td>804</td>
<td>45.2</td>
<td>39.7</td>
<td>5.5</td>
<td>0.299</td>
<td>0.068</td>
</tr>
<tr>
<td>9</td>
<td>510</td>
<td>1843</td>
<td>867</td>
<td>45</td>
<td>39.5</td>
<td>5.5</td>
<td>0.305</td>
<td>0.069</td>
</tr>
<tr>
<td>10</td>
<td>580</td>
<td>2099</td>
<td>737</td>
<td>49</td>
<td>42.3</td>
<td>6.7</td>
<td>0.301</td>
<td>0.049</td>
</tr>
<tr>
<td>11</td>
<td>580</td>
<td>1891</td>
<td>849</td>
<td>47.8</td>
<td>41.5</td>
<td>6.3</td>
<td>0.305</td>
<td>0.052</td>
</tr>
<tr>
<td>12</td>
<td>580</td>
<td>1983</td>
<td>771</td>
<td>50.3</td>
<td>43.9</td>
<td>6.4</td>
<td>0.312</td>
<td>0.054</td>
</tr>
<tr>
<td>13</td>
<td>580</td>
<td>1989</td>
<td>735</td>
<td>48</td>
<td>41.9</td>
<td>6.1</td>
<td>0.321</td>
<td>0.052</td>
</tr>
<tr>
<td>14</td>
<td>850</td>
<td>2072</td>
<td>806</td>
<td>61.2</td>
<td>54.2</td>
<td>7</td>
<td>0.291</td>
<td>0.031</td>
</tr>
<tr>
<td>15</td>
<td>850</td>
<td>2044</td>
<td>918</td>
<td>67.2</td>
<td>60.2</td>
<td>7</td>
<td>0.288</td>
<td>0.028</td>
</tr>
<tr>
<td>16</td>
<td>980</td>
<td>1919</td>
<td>640</td>
<td>63.9</td>
<td>57.5</td>
<td>6.4</td>
<td>0.322</td>
<td>0.078</td>
</tr>
<tr>
<td>17</td>
<td>980</td>
<td>1896</td>
<td>774</td>
<td>64.8</td>
<td>57.1</td>
<td>7.7</td>
<td>0.314</td>
<td>0.058</td>
</tr>
<tr>
<td>18</td>
<td>1100</td>
<td>1923</td>
<td>785</td>
<td>70.4</td>
<td>62.6</td>
<td>7.8</td>
<td>0.310</td>
<td>0.048</td>
</tr>
<tr>
<td>19</td>
<td>1100</td>
<td>1980</td>
<td>732</td>
<td>70.4</td>
<td>62.5</td>
<td>7.9</td>
<td>0.310</td>
<td>0.049</td>
</tr>
<tr>
<td>20</td>
<td>1100</td>
<td>1996</td>
<td>856</td>
<td>72</td>
<td>63.5</td>
<td>8.5</td>
<td>0.288</td>
<td>0.041</td>
</tr>
<tr>
<td>21</td>
<td>1100</td>
<td>2105</td>
<td>819</td>
<td>72.4</td>
<td>63.9</td>
<td>8.5</td>
<td>0.291</td>
<td>0.039</td>
</tr>
</tbody>
</table>

$T_{wz}$ = average travel time in work zone, defined as the average travel time of a vehicle traveling from the start of the transition area to the end of the termination area;

$T_0$ = average travel time outside work zone, defined as the average time taken by a vehicle traveling the distance of $L$ outside work zone;

$D_{wz}$ = average traffic delay per work zone, defined as the difference between travel times in and outside work zone, that is, $D_{wz} = T_{wz} - T_0$

In the HCA model, randomization probability parameters are replaced by randomization probability functions. The randomization probability is modeled as a linear regression function of its influencing factors. Table 3.5 shows the coefficients of independent variables in randomization probability functions. It can be found that
all $p$-values are less than 0.05. It implies that the coefficients in the randomization probability functions are different from zero at the 0.05 level of significance.

<table>
<thead>
<tr>
<th>#</th>
<th>$L_a$ (m)</th>
<th>$L_t$ (m)</th>
<th>$f_L$ (vph)</th>
<th>$f_H$ (vph)</th>
<th>$T_{pw}$ (sec/veh/zone)</th>
<th>$T_0$ (sec/veh/zone)</th>
<th>$D_{pw}$ (sec/veh/zone)</th>
<th>$p_1$</th>
<th>$p_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>40</td>
<td>1322</td>
<td>352</td>
<td>11.42</td>
<td>10.22</td>
<td>1.20</td>
<td>0.350</td>
<td>0.188</td>
</tr>
<tr>
<td>2</td>
<td>110</td>
<td>50</td>
<td>1638</td>
<td>359</td>
<td>14.54</td>
<td>11.94</td>
<td>2.60</td>
<td>0.330</td>
<td>0.168</td>
</tr>
<tr>
<td>3</td>
<td>130</td>
<td>45</td>
<td>1502</td>
<td>286</td>
<td>14.30</td>
<td>12.20</td>
<td>2.10</td>
<td>0.345</td>
<td>0.186</td>
</tr>
<tr>
<td>4</td>
<td>130</td>
<td>45</td>
<td>1661</td>
<td>293</td>
<td>14.48</td>
<td>12.18</td>
<td>2.30</td>
<td>0.336</td>
<td>0.176</td>
</tr>
<tr>
<td>5</td>
<td>170</td>
<td>20</td>
<td>1855</td>
<td>353</td>
<td>14.44</td>
<td>11.64</td>
<td>2.80</td>
<td>0.316</td>
<td>0.138</td>
</tr>
<tr>
<td>6</td>
<td>180</td>
<td>90</td>
<td>738</td>
<td>287</td>
<td>21.12</td>
<td>19.32</td>
<td>1.80</td>
<td>0.417</td>
<td>0.242</td>
</tr>
<tr>
<td>7</td>
<td>180</td>
<td>90</td>
<td>746</td>
<td>276</td>
<td>20.62</td>
<td>19.32</td>
<td>1.30</td>
<td>0.414</td>
<td>0.242</td>
</tr>
<tr>
<td>8</td>
<td>240</td>
<td>65</td>
<td>784</td>
<td>290</td>
<td>22.35</td>
<td>20.05</td>
<td>2.30</td>
<td>0.408</td>
<td>0.228</td>
</tr>
<tr>
<td>9</td>
<td>290</td>
<td>80</td>
<td>717</td>
<td>227</td>
<td>28.68</td>
<td>26.38</td>
<td>2.30</td>
<td>0.421</td>
<td>0.243</td>
</tr>
</tbody>
</table>

Table 3.5 Estimated coefficients associated with the variables for the randomization probability functions

<table>
<thead>
<tr>
<th>Facility</th>
<th>Variable</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressway work zones</td>
<td>Intercept</td>
<td>0.347</td>
<td>19.34</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>$f_L$</td>
<td>-6.6×10^{-5}</td>
<td>-14.57</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>$f_H$</td>
<td>-2.0×10^{-4}</td>
<td>-6.92</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>$L_a$</td>
<td>-1.0×10^{-4}</td>
<td>-5.05</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>$L_t$</td>
<td>2.27×10^{-4}</td>
<td>2.76</td>
<td>0.050</td>
</tr>
</tbody>
</table>

Therefore, the randomization probability functions can be expressed by

(i) Outside work zone
\[ p_1 = \begin{cases} 0.541 - 6.80 \times 10^{-5} f_L - 1.28 \times 10^{-4} f_H, & R^2 = 91.8\% \text{ in expressway} \\ 0.509 - 8.40 \times 10^{-5} f_L - 1.20 \times 10^{-4} f_H, & R^2 = 99.1\% \text{ in arterial road} \end{cases} \] (3.4)

(ii) In work zone

Expressway work zones

\[ p_2 = 0.425 - 1.24 \times 10^{-4} f_L - 1.30 \times 10^{-4} f_H - 3.00 \times 10^{-5} L_t, \quad R^2 = 91.8\% \] (3.5)

Arterial work zones

\[ p_2 = 0.347 - 6.6 \times 10^{-5} f_L - 2.0 \times 10^{-4} f_H - 1.0 \times 10^{-4} L_t + 2.27 \times 10^{-4} L_a, \quad R^2 = 99.1\% \] (3.6)

In Eq. (3.5), the variable \( L_t \) is not contained in the randomization probability function for expressway work zones because \( L_t \) is a constant value of 200 m for expressway work zones according to the Public Works Department’s requirement (PWD, 1985). If there were some expressway work zones with unequal transition area lengths, the variable \( L_t \) would be added into the randomization probability function in expressway work zones.

According to Eqs. (3.4)-(3.6), it can be seen that the randomization probability decreases with the light vehicle flow and heavy vehicle flow. The rationale is that a driver, who is moving at a high speed in the light traffic conditions (e.g., small rates of light/heavy vehicle flow), is not able to focus his/her attention indefinitely (Maerivoet and Moor, 2005). Therefore, the high traveling speed will fluctuate widely in the light traffic conditions. In the heavy traffic conditions (e.g., big rates of light/heavy vehicle flow), the driver is often moving at a low speed and has to focus his/her attention due to the high traffic density. Therefore, drivers will not frequently speed up and slow down when traffic flow increases.
3.4 Model Validation

In this chapter, the model validation is performed at microscopic and macroscopic levels. At the microscopic level, both position and speed of individual vehicles generated from the HCA model are compared against the field data. At the macroscopic level, travel time in and outside of the work zone are respectively compared with the field data.

3.4.1 Microscopic Validation

According to Benekohal (1989), a microscopic validation requires comparing individual vehicle trajectories and speeds from the simulation against the field data. In this chapter, one set of arterial work zone data collected from an arterial work zone located on Ang Mo Kio Avenue 3 of Singapore is used for the microscopic validation. In this arterial work zone, the length of advance warning area $L_w$ is about 140 m. The transition area length $L_t$ and activity area length $L_a$ equal to 30 m and 80 m, respectively. A stretch of 140 m along the advance warning and transition areas is considered. The positions and speeds of 38 individual vehicles including 29 light vehicles and 9 heavy vehicles in a platoon are captured within 75 seconds. In order to generate the same initial headway, the HCA model uses the observed arrival distribution pattern to generate vehicles.
Figure 3.2 Comparison of longitudinal vehicle trajectories in Lane 1 and 2 from the HCA model against the field data

Figure 3.2 (a) and (b) shows the longitudinal trajectories of 38 vehicles in Lane 1 and 2 from the HCA model and the field data. Hereafter, a vehicle initially generated from Lane 1 is defined as a merging vehicle (C) while it is considered to be a through vehicle (V) if it is initially generated from Lane 2. From the figure, it can be clearly seen that the majority of longitudinal positions of the merging/through vehicles from the simulation and from the survey show close agreement for every
second. There exists a slightly large error of the positions of some vehicles (C4, C7, V7 and V8) because of the discrete nature of the model. Nevertheless, the longitudinal trajectory patterns from the simulation are similar to those from the observed data and the errors are acceptable. Similarly, the lateral trajectory patterns as well as lane change durations of merging vehicles are also close to the field data, shown in Figure 3.3 and 3.4.

![Figure 3.3 Comparison of lateral vehicle trajectories of merging vehicles in Lane 1 from the HCA model against the field data](image)

Figure 3.3 Comparison of lateral vehicle trajectories of merging vehicles in Lane 1 from the HCA model against the field data

To further investigate whether the HCA model can describe the dynamic individual behavior well, error tests are also applied to quantitatively measure the closeness of fit of individual vehicle speed from the simulation compared with the field data. In previous studies (e.g., Bham and Benekohal, 2004), the following four error tests are usually used for the comparison of the simulation results with the field
data: 1) Root mean square error ($RMSE$); 2) Root mean square percent error ($RMSPE$); 3) Mean percent error ($MPE$) which indicates the existence of systemic under- or over-prediction in the simulated measurements; and 4) Theil’s inequality coefficient ($U$) that is usually used in econometrics (Pindyck and Rubinfeld, 1998).

\begin{equation}
RMSE = \sqrt{\frac{1}{N_o} \sum_{n=1}^{N_o} (y_n^s - y_n^0)^2}
\end{equation}

where, $y_n^s$ is the $n^{th}$ estimate from the HCA model, $y_n^0$ is the corresponding observed value from the field survey, $N_o$ is the number of observations.

The root mean square percent error ($RMSPE$), mean percent error ($MPE$) and Theil’s inequality coefficient ($U$) are respectively expressed as:

\begin{equation}
RMSPE = \sqrt{\frac{1}{N_o} \sum_{n=1}^{N_o} \left( \frac{y_n^s - y_n^0}{y_n^0} \right)^2}
\end{equation}

\begin{equation}
MPE = \frac{1}{N_o} \sum_{n=1}^{N_o} \left( \frac{y_n^s - y_n^0}{y_n^0} \right)
\end{equation}

Figure 3.4 Comparison of lane change durations of merging vehicles from the HCA model against the field data

The root mean square error ($RMSE$) is defined as:

The root mean square percent error ($RMSPE$), mean percent error ($MPE$) and Theil’s inequality coefficient ($U$) are respectively expressed as:
Chapter 3 Heterogeneous Cellular Automata Model for Work Zone Traffic Delay Estimation

\[
U = \frac{\frac{1}{N_o} \sum_{n=1}^{N_o} (y_n^s - y_n^a)^2}{\sqrt{\frac{1}{N_o} \sum_{n=1}^{N_o} (y_n^s)^2 + \frac{1}{N_o} \sum_{n=1}^{N_o} (y_n^a)^2}}
\]  

(3.10)

The above four tests are performed for the speed of individual vehicles from the simulation versus the field data at each second.

Table 3.6 gives the error test results on individual vehicle speeds. The value of MPE is positive for most of the individual vehicles. Only five vehicles C8, C9, V11, V16 and V21 have negative MPE. The average MPE of all vehicles is positive (≈6%), suggesting that there exists a slight over-estimation of individual speed from the simulation.

The RMSPE is less than 20 percent for 33 vehicles. Only five vehicles including C3, V9, V12, V14 and V17 have RMSPE higher than 20 percent. It should be noted that vehicles C3 and V9 are the two heavy vehicles. The higher RMSPE of these five vehicles may be due to their lower speeds. The smaller denominator (observed speed) of Eq. (3.8) can produce a higher error value. To further investigate these five vehicles, their RMSE and U are then checked. According to Table 3.6, it can be clearly seen that the vehicle C3 has the largest RMSE of 3.97, followed by the vehicle V9 (3.64). Likewise, these two heavy vehicles also have slightly higher U, compared with the other vehicles. One possible reason for the higher RMSE and U of vehicles C3 and V9 could be that there exist measurement errors in the field survey. Nevertheless, the RMSE and U for the rest of vehicles are all acceptable. Meanwhile, the average RMSE is a small value (2.11) and the average U is very close to zero.
### Table 3.6 Error tests for the individual vehicle speed

<table>
<thead>
<tr>
<th>Lane #</th>
<th>Vehicle#</th>
<th>MPE</th>
<th>RMSPE</th>
<th>RMSE</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C1(Light)</td>
<td>6.30%</td>
<td>11.39%</td>
<td>2.15</td>
<td>0.0557</td>
</tr>
<tr>
<td></td>
<td>C2(Light)</td>
<td>14.87%</td>
<td>20.15%</td>
<td>2.87</td>
<td>0.0801</td>
</tr>
<tr>
<td></td>
<td>C3(Heavy)</td>
<td>22.05%</td>
<td>39.65%</td>
<td>3.97</td>
<td>0.1130</td>
</tr>
<tr>
<td></td>
<td>C4(Light)</td>
<td>2.23%</td>
<td>9.90%</td>
<td>1.64</td>
<td>0.0436</td>
</tr>
<tr>
<td></td>
<td>C5(Light)</td>
<td>3.92%</td>
<td>16.97%</td>
<td>3.39</td>
<td>0.0862</td>
</tr>
<tr>
<td></td>
<td>C6(Heavy)</td>
<td>2.15%</td>
<td>4.31%</td>
<td>0.79</td>
<td>0.0210</td>
</tr>
<tr>
<td></td>
<td>C7(Light)</td>
<td>5.20%</td>
<td>12.28%</td>
<td>1.94</td>
<td>0.0526</td>
</tr>
<tr>
<td></td>
<td>C8(Light)</td>
<td>-1.02%</td>
<td>7.27%</td>
<td>1.19</td>
<td>0.0336</td>
</tr>
<tr>
<td></td>
<td>C9(Light)</td>
<td>-2.33%</td>
<td>3.57%</td>
<td>1.00</td>
<td>0.0192</td>
</tr>
<tr>
<td></td>
<td>C10(Heavy)</td>
<td>5.71%</td>
<td>9.38%</td>
<td>1.59</td>
<td>0.0434</td>
</tr>
<tr>
<td></td>
<td>C11(Light)</td>
<td>2.13%</td>
<td>8.41%</td>
<td>1.41</td>
<td>0.0380</td>
</tr>
<tr>
<td></td>
<td>C12(Heavy)</td>
<td>12.58%</td>
<td>18.70%</td>
<td>2.80</td>
<td>0.0785</td>
</tr>
<tr>
<td></td>
<td>V1(Light)</td>
<td>3.73%</td>
<td>8.93%</td>
<td>1.72</td>
<td>0.0464</td>
</tr>
<tr>
<td></td>
<td>V2(Light)</td>
<td>16.39%</td>
<td>19.69%</td>
<td>3.39</td>
<td>0.0930</td>
</tr>
<tr>
<td></td>
<td>V3(Light)</td>
<td>12.80%</td>
<td>15.65%</td>
<td>2.40</td>
<td>0.0679</td>
</tr>
<tr>
<td></td>
<td>V4(Heavy)</td>
<td>11.95%</td>
<td>16.01%</td>
<td>2.47</td>
<td>0.0694</td>
</tr>
<tr>
<td></td>
<td>V5(Light)</td>
<td>0.42%</td>
<td>2.82%</td>
<td>0.53</td>
<td>0.0140</td>
</tr>
<tr>
<td></td>
<td>V6(Light)</td>
<td>1.27%</td>
<td>3.66%</td>
<td>0.67</td>
<td>0.0180</td>
</tr>
<tr>
<td></td>
<td>V7(Light)</td>
<td>3.01%</td>
<td>7.02%</td>
<td>1.22</td>
<td>0.0331</td>
</tr>
<tr>
<td></td>
<td>V8(Light)</td>
<td>4.23%</td>
<td>13.36%</td>
<td>2.09</td>
<td>0.0579</td>
</tr>
<tr>
<td></td>
<td>V9(Heavy)</td>
<td>18.98%</td>
<td>28.13%</td>
<td>3.64</td>
<td>0.1026</td>
</tr>
<tr>
<td></td>
<td>V10(Light)</td>
<td>6.23%</td>
<td>10.59%</td>
<td>1.73</td>
<td>0.0471</td>
</tr>
<tr>
<td></td>
<td>V11(Light)</td>
<td>-1.29%</td>
<td>12.98%</td>
<td>2.29</td>
<td>0.0609</td>
</tr>
<tr>
<td></td>
<td>V12(Light)</td>
<td>16.39%</td>
<td>25.72%</td>
<td>3.40</td>
<td>0.0963</td>
</tr>
<tr>
<td></td>
<td>V13(Light)</td>
<td>7.55%</td>
<td>11.54%</td>
<td>1.64</td>
<td>0.0525</td>
</tr>
<tr>
<td></td>
<td>V14(Light)</td>
<td>14.87%</td>
<td>38.24%</td>
<td>3.13</td>
<td>0.1049</td>
</tr>
<tr>
<td></td>
<td>V15(Light)</td>
<td>2.29%</td>
<td>10.26%</td>
<td>1.87</td>
<td>0.0504</td>
</tr>
<tr>
<td></td>
<td>V16(Light)</td>
<td>-13.14%</td>
<td>19.22%</td>
<td>3.61</td>
<td>0.1032</td>
</tr>
<tr>
<td></td>
<td>V17(Light)</td>
<td>14.10%</td>
<td>24.81%</td>
<td>3.30</td>
<td>0.0914</td>
</tr>
<tr>
<td></td>
<td>V18(Light)</td>
<td>3.39%</td>
<td>9.54%</td>
<td>1.66</td>
<td>0.0447</td>
</tr>
<tr>
<td></td>
<td>V19(Light)</td>
<td>1.14%</td>
<td>13.19%</td>
<td>2.42</td>
<td>0.0647</td>
</tr>
<tr>
<td></td>
<td>V20(Heavy)</td>
<td>6.46%</td>
<td>12.39%</td>
<td>1.96</td>
<td>0.0538</td>
</tr>
<tr>
<td></td>
<td>V21(Light)</td>
<td>-0.67%</td>
<td>7.61%</td>
<td>1.43</td>
<td>0.0374</td>
</tr>
<tr>
<td></td>
<td>V22(Heavy)</td>
<td>6.95%</td>
<td>11.20%</td>
<td>1.85</td>
<td>0.0506</td>
</tr>
<tr>
<td></td>
<td>V23(Light)</td>
<td>0.71%</td>
<td>14.10%</td>
<td>2.43</td>
<td>0.0651</td>
</tr>
<tr>
<td></td>
<td>V24(Light)</td>
<td>8.10%</td>
<td>13.82%</td>
<td>2.07</td>
<td>0.0597</td>
</tr>
<tr>
<td></td>
<td>V25(Heavy)</td>
<td>2.44%</td>
<td>4.87%</td>
<td>0.88</td>
<td>0.0239</td>
</tr>
<tr>
<td></td>
<td>V26(Light)</td>
<td>8.87%</td>
<td>12.28%</td>
<td>2.00</td>
<td>0.0559</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>6.08%</td>
<td>13.94%</td>
<td>2.11</td>
<td>0.0587</td>
</tr>
</tbody>
</table>

*Vehicle type in parentheses*
The relatively small and acceptable errors of the individual vehicle speed present strong evidence that the HCA model can well describe work zone traffic dynamics at the microscopic level.

### 3.4.2 Macroscopic Validation

To further investigate the model validity, a macroscopic evaluation is also performed to identify the overall model performance. Actual travel time of each vehicle can be easily measured by observing the vehicle arrival times in two license plates, which are placed in two positions. Therefore, in macroscopic validation, the average travel times in and outside work zone are respectively used to evaluate how well the HCA model performs. In this section, another 20 PIE expressway work zone data sets and 6 arterial work zone data sets are used for the macroscopic validation.

Figure 3.5(a) shows a comparison of the average travel time from the HCA simulation against the observed average travel time in and outside expressway work zone. The solid line represents \( T = T \), suggesting that the estimated travel time \( \hat{T} \) fully matches the observed value \( T \). The slight departure of the scattered dots indicates that the average travel time from the HCA model is perfectly close to the measured value for each expressway work zone data. The comparison results from the simulation and from the observed data in Figure 3.5(b) also show a very good agreement for the arterial work zones. These satisfactory simulated results imply that the HCA model is able to simulate the real-world work zone traffic with a high degree of accuracy at the macroscopic level.
Figure 3.5 Comparison of average travel time from the HCA simulation versus the field data
3.5 Model Application

As mentioned earlier, the accurate estimation of traffic delay is of utmost importance for traffic engineers because it can affect the efficiency of work zone plans and traffic management strategies. Since the HCA model can simulate work zone traffic well, it is theoretically able to provide an accurate estimate of traffic delay occurred in work zone.

3.5.1 Traffic Delay Estimation

Figure 3.6 (a) and (b) shows the estimated traffic delays from the HCA model on 20 expressway work zones and 6 arterial work zones, which have been used for the macroscopic validation. It can be seen that the estimated traffic delay for both work zone types closely match the observed data. The figure visually confirms that the HCA model could provide an accurate estimate of work zone traffic delay.

To support the visual claim, the error tests are also carried out for the traffic delay, and the corresponding results are shown in Table 3.7. It can be seen the statistic errors are very small. For the accuracy comparison, the microscopic simulation tool—PARAMICS is applied to estimate work zone traffic delay as well. In PARAMICS, the three sensitive system parameters including mean target headway, mean reaction time and minimum gap are calibrated using a conventional genetic algorithm. The calibrated values for the three parameters are respectively 1.5 sec, 0.6 sec and 1.5 m for the expressway work zones. For the arterial work zones, the calibrated values are 1.25 sec, 1.0 sec and 1.0 m for the three parameters, respectively.
Figure 3.6 Comparisons of traffic delay from the HCA simulation against the field data

Table 3.7 also shows the statistical simulated results from PARAMICS. It can be found that the estimation accuracy of traffic delay, for the expressway work zones, from PARAMICS is a little lower than that from the HCA model. However, the HCA model provides much higher accuracy than PARAMICS for the estimation of traffic delay for the arterial work zones.

To evaluate the computational performance of the HCA model, vehicles ranging 500 to 3000 with increments of 500 vehicles are simulated for 720 sec. The heavy vehicle percentage is assumed to be 10%. An expressway with the length of
4000 m is simulated comprising of an entrance and an exit. The total work zone length (i.e., $L_w$) is assumed to be 730 m. The simulation is implemented on a Dell Optiplex computer with an Intel Core 2 CPU (2.40 GHz) and a RAM of 1.99 GB.

Table 3.8 shows the execution times for the HCA model and PARAMICS. From this table, it can be found that the HCA model is much faster than PARAMICS at different traffic conditions. Therefore, it can be concluded that the HCA model is a good alternative for work zone traffic delay estimation in terms of accuracy.

### Table 3.7 Statistical simulated traffic delay results from the proposed HCA model and PARAMICS

<table>
<thead>
<tr>
<th>Model</th>
<th>Work Zone Type</th>
<th>$MPE$</th>
<th>$RMSPE$</th>
<th>$U$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCA model</td>
<td>Expressway work zone</td>
<td>2.21%</td>
<td>9.38%</td>
<td>0.0456</td>
<td>0.9039</td>
</tr>
<tr>
<td></td>
<td>Arterial work zone</td>
<td>5.78%</td>
<td>7.82%</td>
<td>0.0359</td>
<td>0.9157</td>
</tr>
<tr>
<td>PARAMICS</td>
<td>Expressway work zone</td>
<td>-4.11%</td>
<td>13.74%</td>
<td>0.0721</td>
<td>0.8348</td>
</tr>
<tr>
<td></td>
<td>Arterial work zone</td>
<td>11.74%</td>
<td>24.82%</td>
<td>0.0824</td>
<td>0.5533</td>
</tr>
</tbody>
</table>

### Table 3.8 Execution times for the HCA model and PARAMICS

<table>
<thead>
<tr>
<th>Traffic flow (vph)</th>
<th>Simulation runs</th>
<th>Real traffic simulated</th>
<th>Execution time (i.e., CPU time)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>HCA model</td>
</tr>
<tr>
<td>500</td>
<td>28</td>
<td>720 sec</td>
<td>2.3 sec</td>
</tr>
<tr>
<td>1000</td>
<td>28</td>
<td>720 sec</td>
<td>5.4 sec</td>
</tr>
<tr>
<td>1500</td>
<td>28</td>
<td>720 sec</td>
<td>10.1 sec</td>
</tr>
<tr>
<td>2000</td>
<td>28</td>
<td>720 sec</td>
<td>16.2 sec</td>
</tr>
<tr>
<td>2500</td>
<td>28</td>
<td>720 sec</td>
<td>20.3 sec</td>
</tr>
<tr>
<td>3000</td>
<td>28</td>
<td>720 sec</td>
<td>26.9 sec</td>
</tr>
</tbody>
</table>
3.5.2 Sensitivity Analysis of Traffic Delay

Since the HCA model can provide more accurate estimates of work zone traffic delay than PARAMICS, it is further employed to examine the effects of influencing factors on work zone traffic delay. For simplicity, the variables of light vehicle flow $f_L$ and heavy vehicle flow $f_H$ are converted to the variables of traffic flow $f$ and heavy vehicle percentage $hv$, where

$$f = f_L + f_H \quad (3.11)$$

$$hv = \frac{f_H}{f_L + f_H} \quad (3.12)$$

In addition to the traffic delay index $D_{wz}$, another index $d_{wz}$, in terms of sec veh/100m, is also used to express the work zone traffic delay. In order to assess the impacts of the transition area length $L_t$ and the activity area length $L_a$, these two parameters are varied.
Figure 3.7 Effects of the transition area length and the activity area length on $D_{wz}$

Figure 3.7 (a) and (b) graphically show the effects of $L_t$ and $L_a$ on traffic delay occurred in the whole work zone ($D_{wz}$). Unless used as a variable in the figure, the parameters $L_t, L_a, f$ and $hv$ are respectively taken as 20 m, 90 m, 1700 vph and 20% to plot the curves in the figure. It can be clearly seen from this figure that $D_{wz}$ monotonically increases with the transition area length $L_t$ and the activity area length $L_a$.
Comparing with Figure 3.7 (a) and (b), it can be further found that \( L_t \) has a much bigger effect than \( L_a \) on \( D_{wz} \), especially in the light traffic condition. Take the traffic flow of 800 vph as an example, the increase of 50 m of \( L_t \) can cause an increase of 1.35 sec/veh/zone while the increase of the same length of \( L_a \) can only cause an increase of 0.15 sec/veh/zone of the average traffic delay.

**Figure 3.8 Effects of the transition area length and the activity area length on \( d_{wz} \)**
Figure 3.8 (a) and (b) graphically depicts the effects of $L_t$ and $L_a$ on work zone traffic delay per 100m ($d_{wz}$). Similar to $D_{wz}$, $d_{wz}$ increases with the transition area length $L_t$, shown in Figure 3.8 (a). One possible reason is that the randomization probability parameter $p_2$ increases as $L_t$ increases. The high value of randomization probability could cause a lower traveling speed, further resulting in larger traffic delay. However, the randomization probability parameter $p_2$ decreases with $L_a$ while the smaller $p_2$ will reduce $d_{wz}$. This is the major reason why $d_{wz}$ does not monotonically increase with the activity area length $L_a$, shown in Figure 3.8 (b).

Figure 3.9 Impacts of traffic flow and heavy vehicle percentage on $D_{wz}$

Figure 3.9 shows the impacts of $f$ and $hv$ on $D_{wz}$. Similar to Figure 3.7 and 3.8, Figure 3.9 clearly shows that $D_{wz}$ increases with traffic flow $f$. It can be also seen from this figure that $D_{wz}$ increases as the heavy vehicle percentage $hv$ increases.

Since $D_{wz}$ is sensitive to the factors of $L_t$, $L_a$, $f$ and $hv$, it can be expressed as a
function of the four factors with the following functional form:

\[ D_{wz} = \alpha_0 + \sum_{i=1}^{N_a} (\alpha_i \times x_i) + \epsilon_1 \]  

(3.13)

where \( \alpha_0, \alpha_i, i = 1, \ldots, N_a \) are the coefficients to be estimated; \( x_i \) are the explanatory variables associated with the above four factors variables, \( N_a \) is the number of explanatory variables.

It is assumed that each of the explanatory variables has four possible relations with respect to its relevant contributing factor: identical, logarithm, square and combination. For example, the explanatory variable \( x_1 \) related to the transition area length can be expressed by \( L_t, \ln(L_t), L_t^2, L_t \cdot f, L_t \cdot hv, L_t^2 \cdot f \) and \( L_t^2 \cdot hv \). Therefore, the maximum number of explanatory variables is \( 2 \times 7 + 9 = 23 \) for Eq. (3.13).

The backward elimination method, in which the variables are tested for the removal from the function one by one based on the significant level of the likelihood ratio, is used to select the variables for the work zone traffic delay function shown by Eq. (3.13). Initially, all 23 explanatory variables are selected by the function \( D_{wz} \). We repeatedly remove the least significant variable with the level of significance greater than 0.05. After each removal, variables that are still kept in the function will be tested for next possible removal until all the variables kept in the function are statistically significant at the level of 0.05.

Based on 78 sets of traffic delay data, the function \( D_{wz} \) can be calibrated using the above backward elimination method. The results are shown as follows:
\begin{equation}
D_{wz} = 55.25 - 4.97 \ln(f) + 3.68 \times 10^{-3} f - 1.52 \times 10^{-2} L_a + 2.38 \times 10^{-5} f \cdot L_a \\
+ 2.76 \times 10^{-5} f \cdot L_t + 15.73 \ln(hv), \quad R^2 = 98.7\%
\end{equation}

The function $D_{wz}$ can be further applied to estimate the marginal effects of the activity area length, the transition area length, the traffic flow and the heavy vehicle percentage. For instance, it is assumed that there is a work zone with the following configuration: $L_t = 20$ m, $L_a = 90$ m, $f = 1700$ vph and $hv = 20\%$. According to Eq. (3.14), one unit increase of $L_a$ will result in an increase of 0.025 sec/veh/zone while one unit increase of $L_t$ will increase traffic delay by 0.047 sec/veh/zone. In this case, the marginal effect of $L_t$ is about twice ($\approx 0.047/0.025$) that of $L_a$. Compared with the transition area and activity area lengths, traffic flow $f$ has a much lower marginal effect ($\approx 0.0035$ sec/veh/zone).

### 3.6 Summary

Due to its simplicity and high computational efficiency, the CA model has been used to simulate road traffic in the past. This chapter has developed a HCA model to simulate heterogeneous work zone traffic as well as estimate work zone traffic delay. A fine cell size is decided for the HCA model such that it can give a better representation of the unit of vehicle in the heterogeneous work zone traffic. The fine cell size is taken as 0.5m (length) $\times$ 0.7m (width). The vehicle updating procedure includes two parts. The first part updates the vehicles’ longitudinal speeds and positions using the improved forwarding rules, while the second part updates the vehicles’ lateral speeds and positions using more realistic lane changing rules. Since
driver behavior at work zones can be influenced by the factors of work zone configuration and traffic flow, the randomization probability parameter in the HCA model is formulated as a linear function of the influencing factors.

The developed HCA model has been calibrated and validated both microscopically and macroscopically in detail using real expressway and arterial work zone data collected from Singapore. The graphical and error tests are conducted to test the validity of the model. From the validation results, it can be found that the model simulates the heterogeneous traffic in the work zone well. After validation, the model is then applied to estimate traffic delay occurred in work zone. The results show that the HCA model outperforms PARAMICS in estimating traffic delay in terms of accuracy, suggesting that the HCA model is a good alternative for work zone traffic delay estimation. Finally, a sensitivity analysis is carried out using the HCA model to examine the marginal effects of the activity area length, the transition area length, traffic flow and heavy vehicle percentage on work zone traffic delay.

Since the HCA model could provide high estimation accuracy of traffic delay, it will be incorporated by a total work zone cost minimization model which is developed to determine the optimal subwork zone operational strategy in Chapter 5. In addition, the deterministic queuing model is usually incorporated in the total work zone cost minimization model. The accurate estimation of work zone capacity is very important because it is a key input for the deterministic queuing model. In next chapter, a decision tree approach will be used to accurately estimate work zone capacity for the deterministic queuing model.
CHAPTER 4 DECISION TREE-BASED MODEL FOR WORK ZONE CAPACITY ESTIMATION

4.1 Introduction

As mentioned in Chapter 3, the HCA model is a good alternative to estimate traffic delay because of its high estimation. However, it is not widely used by traffic engineers because of its complexity. The deterministic queuing model is widely employed by traffic engineers for traffic delay estimation to determine an optimal subwork zone operational strategy because of its simplicity. Since work zone capacity is a key parameter of the deterministic queuing model, the ability to estimate work zone capacity accurately is imperative for traffic engineers when the deterministic queuing model is used to estimate work zone traffic delay.

Various studies (Krammes and Lopez, 1994; Kim et al., 2001; Al-Kaisy and Hall, 2003; Sarasua et al., 2006) have been conducted to estimate work zone capacity and numerous factors including work zone duration, heavy vehicle percentage, work time, intensity of work activity, weather conditions, lane closure configuration, road type and grade have been found to influence work zone capacity. However, not all the important influencing factors are incorporated by any one capacity estimation model proposed in previous studies. The lack of important influencing factors may lead to low estimation accuracy. Take the work zone capacity estimation model addressed by Krammes and Lopez (1994) as an example, their model only takes the factors of work intensity, ramp, heavy vehicle percentage and the number of lanes into consideration, whereas the factors such as work time, weather and road type are neglected. Due to
the neglect of the factor of work time, the estimate of work zone capacity at night is equal to that at daytime while nighttime work zones actually have lower capacities than daytime work zones. In addition, their model is developed based on the short-term work zone data, which may greatly underestimate the capacities of long-term work zones.

Since a large number of factors should be taken into account, it would be inappropriate to estimate work zone capacity using one simple model. The current HCM (2000) provides two distinct work zone capacity estimation guidelines, applicable to the short-term work zones and long-term work zones, respectively. Nevertheless, the HCM also requires traffic engineers to decide on the adjustment factors based on their prior experience, which also leads to significant estimated error due to subjective judgment. Similar to the existing capacity estimation models, the HCM also excludes some important factors, such as lane closure location and road type, which may lead to less accurate estimates. Adeli and Jiang (2003) addressed a neural-fuzzy logic model incorporating 17 influencing factors to estimate work zone capacity. Their model exhibits a higher degree of estimation accuracy, compared with Krammes and Lopez’s (1994) model and the model of Kim et al. (2001). However, it is so complex that it has poor applicability.

Since a variety of factors can affect work zone capacity, there will be a number of work zone scenarios with distinct work zone capacity. To determine all possible work zone scenarios and the corresponding work zone capacity, the data can be partitioned into a number of groups and fit a simple model in each group. A decision tree method is a particularly efficient method for partitioning data. Not surprisingly, decision trees can give better predictions for work zone capacity because trees are well suited for analyzing the complex data that include a large number of variables.
Chapter 4 Decision Tree-based Model for Work Zone Capacity Estimation

Therefore, the decision tree method is used to estimate work zone capacity in this chapter. Note that the decision tree method has already been applied in various areas including the medicine, business and industry (Abrahams et al., 2009) because of its high prediction accuracy and ease of use.

4.2 Model Formulation

4.2.1 Factors Affecting Work Zone Capacity

As mentioned in Chapter 2, a large number of variables have been found to affect work zone capacity, including:

1. Heavy vehicle percentage ($x_1$). Note that heavy vehicles include the construction vehicles entering and leaving the work zone.
2. Work zone grade ($x_2$).
3. Work intensity ($x_3$).
4. Road type (rural or urban) ($x_4$).
5. Number of open lanes ($x_5$).
6. Number of closed lanes ($x_6$).
7. Work zone duration (short-term or long-term) ($x_7$).
8. Work time (day or night) ($x_8$).
9. Lane width ($x_9$).
10. Lane closure location (left or right) ($x_{10}$).
11. Work zone length ($x_{11}$).
12. Weather condition (clear, rain) ($x_{12}$).
13. Driver composition (0 or 1) ($x_{13}$).
14. Ramp (yes or no) ($x_{14}$).
Chapter 4 Decision Tree-based Model for Work Zone Capacity Estimation

(15) Work zone speed \( (x_{15}) \).

(16) State/city \( (x_{16}) \).

Therefore, work zone capacity, denoted by \( y \), can symbolically be expressed as a function of these 16 variables, shown as follows:

\[
y = f(x_1, x_2, \ldots, x_{16})
\]

(4.1)

The work zone capacity cannot be described by any simple mathematical function because of a large number of interacting variables. In this chapter, a decision tree is used to describe the complex relationship between work zone capacity and its influencing variables.

### 4.2.2 Decision Tree

A decision tree is a flow-chart-like tree structure where the root node is at the top and the leaf nodes are at the bottom. When the target variable is categorical, the decision tree is called a classification tree, whereas it is a regression tree as the target variable is continuous. Since the work zone capacity is a continuous target variable, the decision tree is actually a regression tree in this chapter.

In a decision tree, the root node contains the entire dataset and the tree grows through the test of partitioning data at the nodes. The outgoing branches of a node correspond to all the possible outcomes of the test at the node. The leaves indicate the groups (scenarios). The point \((x_1, x_2, \ldots, x_{16})\) belongs to a leaf if it falls in the corresponding group. To figure out which group of a data point belongs to, we can start at the root node of the tree and trace a path down the tree according to features of the data point. In addition, each leaf will be assigned by a simple model to estimate work zone capacity. For simplicity, this chapter assumes that the estimated work zone capacity from the simple model is equal to the sample mean of data points in that leaf.
Chapter 4 Decision Tree-based Model for Work Zone Capacity Estimation

More specifically, suppose that a sample of \( c \) data points belong to the leaf \( l \), the simple model for this leaf can be expressed as 
\[
\mu = \frac{1}{N} \sum_{i=1}^{N} y_i,
\]
where \( \mu \) is the estimated work zone capacity and \( y_i \) is the observed work zone capacity in the leaf \( l \).

4.2.3 Decision Tree Construction

The construction of a decision tree generally consists of three steps: tree growing, tree pruning and rule extracting. Training data are applied to grow a decision tree in the first step. Due to the possible over-fitting problem, checking data are used to prune the grown decision tree in the second step. These steps are graphically depicted in Figure 4.1 and in detail illustrated in the following subsections.

Step 1: Tree growing

The principle behind growing a decision tree is to recursively partition the target variable so that the data in descendent nodes are always purer than the data in the parent node. When a training dataset enters the root node of a decision tree, a test is performed to search for all possible splits for all variables using a splitting criterion, which measures the quality of each possible split. Various splitting criterion can be used to grow the decision tree, such as the variance reduction and \( F \)-test splitting criterion. Since work zone capacity is a continuous target variable, this chapter thus uses the \( F \)-test splitting criterion. The following \( F \)-test splitting procedure is implemented to select which variable and split scheme can be used to best split a parent node:

(i) For a given parent node \( t \) at the bottom, the best split \( S'(X, t) \) of a specific variable \( X \) is selected among all possible splits according to the case values of variable \( X \).
Here, the best split refers to the split that has the smallest $p$-value based on the $F$-test. In this chapter, the smallest $p$-value implies that the means of work zone capacity are most significantly different for different categories split by the variable $X$. $p$-value can be calculated from the $F$-statistic:

$$F = \frac{\sum_{i=1}^{N_s} \sum_{n=1}^{N_f} I(x_n = i)(\bar{y}_i - \bar{y})^2 / (N_s - 1)}{\sum_{i=1}^{N_s} \sum_{n=1}^{N_f} I(x_n = i)(y_n - \bar{y}_i)^2 / (N_f - N_s)}$$

(4.2)

$$p\text{-value} = \Pr\left(F(N_s - 1, N_f - N_s) > F\right)$$

(4.3)

where

$$\bar{y}_i = \frac{\sum_{n=1}^{N_f} y_n I(x_n = i)}{\sum_{n=1}^{N_f} I(x_n = i)}$$

(4.4)

$$I(x_n = i) = \begin{cases} 1, & \text{if the case value of variable } X \text{ in the } n^{th} \text{ data belongs to the } i^{th} \text{ category} \\ 0, & \text{otherwise} \end{cases}$$

(4.5)

where $y_n$ is the observed work zone capacity of the $n^{th}$ data set, $N_s$ is the number of categories of variable $X$, $N_f$ is the number of datasets.

For a continuous variable, such as the heavy vehicle percentage, the datasets in the node $t$ are first sorted according to the case value of that continuous variable in an ascending order. Second, the sorted datasets are divided into a finite number of groups with distinct interval lengths and boundaries to create categories. Using the enumeration method, the best interval length and boundary of each group can be determined. The corresponding split is considered as the best split of the continuous variable, $S^*(X, t)$.

(ii) By repeating the sub-step (i), the best splits for all the variables can be determined. Among these best splits, the best discriminating variable $X^*$ which has
the global best split $S^*(X^*, t)$ that has the global smallest $p$-value can be selected.

(iii) If the global smallest $p$-value of the sub-step (ii) is less than or equal to a predetermined $\alpha_{\text{split}}$, the parent node $t$ is split using the global best split $S^*(X^*, t)$ and then go to the sub-step (i). Otherwise, the parent node $t$ is not split and regarded as a leaf node.

(iv) Splitting will be stopped if one of the following two stopping splitting rules is satisfied: a) any node at the bottom cannot be further split because $F$-test is insignificant for the node or the number of datasets in the node is less than the preset minimum number; b) the current tree depth reaches the preset maximum tree depth. Otherwise, go to the sub-step (i).

**Step 2: Tree pruning**

Overly large trees could result in higher prediction mean square error when applied to estimate new datasets (Breiman et al., 1984). In this case, there is a critical need to remove a part of branches that do not contribute to the prediction accuracy of new datasets. There are two different pruning techniques to decision tree pruning. One is pre-pruning approach and the other is post-pruning approach that removes some branches after the tree has been grown. Conceptually, the process of post-pruning tree is analogous to backward model selection procedure commonly used in regression analysis. In this chapter, a post-pruning approach that aims to minimize the prediction mean square for the checking data (Quinlan, 1993) is employed to prune the tree. Using the checking data, those branches that can produce large mean square error are repeatedly removed until the removal of branches cannot contribute to the decrease of mean square error.
Figure 4.1 Three steps for the development of a decision tree

Step 3: Rule extracting

To better understand the tree-based model, the knowledge in the decision tree should be extracted. In the third step, the rules are extracted from the prune tree. One
rule can be created for each path from the root node to a leaf node in the form of an “if-then” rule. All variable values along a path form a conjunction in the rule antecedent (the “if” part), whereas the leaf node determines the predicted mean value of target variable covered by the rule antecedent, forming the rule consequent (the “then” part). The extracted “if-then” rules are easier for humans to understand and can further be incorporated into an expert system.

Note that there may be some redundant rules if the tree size is large. Therefore, any condition in the rule antecedent that does not improve the performance of the rule should be removed. For instance, consider the following rule “If $x_1<10$ and $x_2= (2$ or $3)$ and $x_2=3$ and $x_1<5$, then $y=1$”. This rule can be reduced to “If $x_1<5$ and $x_2=3$, then $y=1$”.

4.3 Data Collection

The data for training, checking and evaluating the decision tree-based model are collected from the existing literature and work zone projects. A total of 182 datasets from fourteen states and cities of U.S. are collected, including 12 sets from Maryland (Kim et al., 2001), 72 sets from Texas (Dudek and Richards, 1981; Krammes and Lopez, 1994), 5 sets from North Carolina (Dixon et al., 1996), 17 sets from California (Krammes and Lopez, 1994), 12 sets from Indiana (Jiang, 1999), 4 sets from Toronto (Ahmed et al., 2000), 19 sets from South Carolina (Sarasua et al., 2006), 21 sets from Florida (Heaslip et al., 2008; Heaslip et al., 2009) and 20 sets from other cities (Borchard et al., 2009). The hourly traffic volume under congested traffic conditions (Texas). Since a majority of work zones take the hourly traffic volume under congested traffic conditions as the work zone capacity, work zone capacity is thus defined as the hourly traffic volume under congested traffic conditions.
in this chapter. It should be noted that only a few datasets provide all the input values of sixteen variables in the decision tree-based capacity estimation model. According to the collected datasets, it can be found that the minimum number of variables provided is equal to five (number of open lanes, number of closed lanes, work intensity, work zone duration and state/city). For a work zone dataset, the unavailable variables are assigned by “unknown” values.

Work zone lengths are quantitatively measured for some work zone datasets (from Maryland) while some are simply categorized into three types: short, medium and long in San Antonio. For the sake of easier presentation, work zone lengths from Maryland are classified into the types of short, medium and long. Here, the short work zone length represents the length less than 1.0 km and the medium work zone length is regarded as the length less than 2.0 km but larger than 1.0 km. The work zone length not less than 2.0 km is considered to be long. Work intensity can be quantitatively determined using the work intensity ratio function proposed by Benekohal et al. (2003). However, it is difficult to obtain the information regarding the number of workers, the number of large construction equipments and the lateral distance for the work intensity ratio function. In general, work intensity is a qualitative and subjective concept without any standard classification scheme. In previous studies, most researchers categorized work intensity according to the type of work activity. In this chapter, it is divided into three levels: 1) Light level for the activity types of barrier/guardrail installation/repair and pavement repair; 2) Medium level for the resurfacing/asphalt removal and stripping removal; and 3) Heavy level for the pavement marking and bridge repair.

The collected 182 work zone datasets summarized in Table 4.1 are further divided into three groups: one for training, another for checking and the third group is
used for evaluation. More specifically, 146 sets of work zone datasets are randomly selected for tree training and 18 sets are for checking the decision tree. The remaining 18 work zone datasets are used for evaluating the decision tree.

### Table 4.1 Collected work zone datasets for training, checking and evaluating decision tree-based capacity estimation model

<table>
<thead>
<tr>
<th>State/City (x₁₆)</th>
<th>Value</th>
<th>182 datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Training</td>
</tr>
<tr>
<td>Maryland</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>North Carolina</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Texas</td>
<td>3</td>
<td>58</td>
</tr>
<tr>
<td>California</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Indiana</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Toronto</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>South Carolina</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>Florida</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Houston</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Orange</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Pearland</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>San Antonio</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>146</strong></td>
</tr>
</tbody>
</table>

#### 4.4 Model Training and Checking

After trying several different values, a significance level of 0.10 for the $F$-test splitting criterion, the minimum number of five datasets in a leaf and the maximum tree depth of 16 are used for the proposed decision tree-based model. The learning process of the decision tree-based model based on 146 training datasets and 18 checking datasets is depicted in Figure 4.2. According to the figure, it can be seen that the prediction mean square error between the predicted work zone capacity and the observed value for the training data monotonically decreases with the number of leaves. However, it is found that the increase of the number of leaves cannot contribute to the decrease of mean square error for the checking data, when the
number of leaves is larger than 16. Therefore, the pruned decision tree comprising 16 leaves (leaf nodes) is considered as the optimal decision tree for work zone capacity estimation, as shown in Figure 4.3. Note that the rounded rectangles in Figure 4.3 represent the leaf nodes of decision tree. Each node consists of the information regarding the number of training datasets ($N$), the values of predicted work zone capacity ($\mu$) and the corresponding standard deviation ($\sigma$).

![Figure 4.2 Learning process of the decision tree model for work zone capacity estimation](image)

As mentioned earlier, a decision tree should be extracted into a set of “if-then” rules so that traffic engineers can quickly apply them to estimate work zone capacity. Table 4.2 lists a total of 16 “if-then” rules extracted from the optimal decision tree. The interpretation of “if-then” rules is straightforward. Traffic engineers can easily estimate work zone capacity by searching for an “if-then” rule with respect to a given work zone. If traffic engineers are not able to find an “if-then” rule for a given work zone with missing data, they can also estimate work zone capacity by tracing a path from the tree down to a node in the sub-tree they do reach.
Figure 4.3 The optimal decision tree for work zone capacity estimation

- : Leaf node

\( N \): Number of datasets in the node

\( \mu \): Predicted average work zone capacity

\( \sigma \): Standard deviation of predicted work zone capacity

= Short-term

\( N: 140 \)

\( \mu: 1519 \)

\( \sigma: 246 \)

Work zone duration

= Long-term

\( N: 36 \)

\( \mu: 1803 \)

\( \sigma: 197 \)

Number of opened lanes
### Table 4.2 Extracted “If-then” rules from the optimal decision tree

<table>
<thead>
<tr>
<th>Rule No.</th>
<th>Rule Antecedent (“If”)</th>
<th>Rule Consequent (“Then”)</th>
<th>Mean capacity (vphpl)</th>
<th>Std. capacity (vphpl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>If road type = “Rural”, work intensity = “Light”, work zone duration = “short-term”</td>
<td></td>
<td>1392</td>
<td>143</td>
</tr>
<tr>
<td>2</td>
<td>If lane closure location = “Left”, work intensity = “Heavy”, work zone duration = “short-term”</td>
<td></td>
<td>1305</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>If lane closure location = “Right”, work intensity = “Heavy”, work zone duration = “short-term”</td>
<td></td>
<td>1138</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>If heavy vehicle percentage &lt; 4.3%, work intensity = “Unknown”, work zone duration = “short-term”</td>
<td></td>
<td>1614</td>
<td>117</td>
</tr>
<tr>
<td>5</td>
<td>If heavy vehicle percentage &gt;= 4.3%, work intensity = “Unknown”, work zone duration = “short-term”</td>
<td></td>
<td>1449</td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>If heavy vehicle percentage &lt; 15.6%, number of open lanes = “1”, work zone duration = “long-term”</td>
<td></td>
<td>1793</td>
<td>129</td>
</tr>
<tr>
<td>7</td>
<td>If heavy vehicle percentage &gt;= 15.6%, number of open lanes = “1”, work zone duration = “long-term”</td>
<td></td>
<td>1616</td>
<td>105</td>
</tr>
<tr>
<td>8</td>
<td>If weather = “clear”, number of open lanes &gt;= 2”, work zone duration = “long-term”</td>
<td></td>
<td>2002</td>
<td>55</td>
</tr>
<tr>
<td>9</td>
<td>If weather = “rainy”, number of open lanes &gt;= 2”, work zone duration = “long-term”</td>
<td></td>
<td>1672</td>
<td>148</td>
</tr>
<tr>
<td>10</td>
<td>If work time = “Night”, road type = “Urban”, work intensity = “Light”, work zone duration = “short-term”</td>
<td></td>
<td>1510</td>
<td>59</td>
</tr>
<tr>
<td>11</td>
<td>If work time = “Day”, road type = “Urban”, work intensity = “Light”, work zone duration = “short-term”</td>
<td></td>
<td>1704</td>
<td>48</td>
</tr>
<tr>
<td>12</td>
<td>If number of closed lanes = “1”, number of open lanes = “1”, work intensity = “Medium”, work zone duration = “short-term”</td>
<td></td>
<td>1383</td>
<td>138</td>
</tr>
<tr>
<td>13</td>
<td>If number of closed lanes &gt;= 2”, number of open lanes = “1”, work intensity = “Medium”, work zone duration = “short-term”</td>
<td></td>
<td>1113</td>
<td>52</td>
</tr>
<tr>
<td>14</td>
<td>If number of closed lanes = “1”, number of open lanes &gt;= 2”, work intensity = “Medium”, work zone duration = “short-term”</td>
<td></td>
<td>1536</td>
<td>82</td>
</tr>
<tr>
<td>15</td>
<td>If number of closed lanes = “2”, number of open lanes &gt;= 2”, work intensity = “Medium”, work zone duration = “short-term”</td>
<td></td>
<td>1454</td>
<td>83</td>
</tr>
<tr>
<td>16</td>
<td>If number of closed lanes &gt;= 3”, number of open lanes &gt;= 2”, work intensity = “Medium”, work zone duration = “short-term”</td>
<td></td>
<td>1371</td>
<td>38</td>
</tr>
</tbody>
</table>
Table 4.3 Values for 18 work zone scenarios for evaluating the decision tree-based model

<table>
<thead>
<tr>
<th>Site</th>
<th>Heavy vehicle percentage</th>
<th>Work zone grade</th>
<th>Work intensity</th>
<th>Road type</th>
<th>No. of open lanes</th>
<th>No. of closed lanes</th>
<th>Work zone duration</th>
<th>Work time</th>
<th>Lane width (m)</th>
<th>Lane closure location</th>
<th>Work zone length</th>
<th>Weather</th>
<th>Driver composition</th>
<th>Ramp</th>
<th>Speed</th>
<th>State/city</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.0%</td>
<td>0</td>
<td>Medium</td>
<td>Urban</td>
<td>2</td>
<td>2</td>
<td>Short-term</td>
<td>Night</td>
<td>3.6</td>
<td>Left</td>
<td>Long</td>
<td>Clear</td>
<td>0</td>
<td>Yes</td>
<td>22</td>
<td>Maryland</td>
</tr>
<tr>
<td>2</td>
<td>3.2%</td>
<td>–</td>
<td>–</td>
<td>Urban</td>
<td>1</td>
<td>1</td>
<td>Short-term</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Left</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Texas</td>
</tr>
<tr>
<td>3</td>
<td>13.2%</td>
<td>–</td>
<td>–</td>
<td>Urban</td>
<td>1</td>
<td>1</td>
<td>Short-term</td>
<td>–</td>
<td>–</td>
<td>Right</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Texas</td>
</tr>
<tr>
<td>4</td>
<td>8.5%</td>
<td>–</td>
<td>–</td>
<td>Urban</td>
<td>2</td>
<td>2</td>
<td>Short-term</td>
<td>–</td>
<td>–</td>
<td>Left</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Texas</td>
</tr>
<tr>
<td>5</td>
<td>3.7%</td>
<td>–</td>
<td>–</td>
<td>Urban</td>
<td>3</td>
<td>1</td>
<td>Short-term</td>
<td>–</td>
<td>–</td>
<td>Left</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Texas</td>
</tr>
<tr>
<td>6</td>
<td>–</td>
<td>Medium</td>
<td>Urban</td>
<td>1</td>
<td>2</td>
<td>Short-term</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Left</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Texas</td>
</tr>
<tr>
<td>7</td>
<td>–</td>
<td>Light</td>
<td>Urban</td>
<td>2</td>
<td>1</td>
<td>Short-term</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Left</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Texas</td>
</tr>
<tr>
<td>8</td>
<td>–</td>
<td>Medium</td>
<td>Urban</td>
<td>3</td>
<td>1</td>
<td>Short-term</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Left</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Texas</td>
</tr>
<tr>
<td>9</td>
<td>–</td>
<td>Heavy</td>
<td>–</td>
<td>3</td>
<td>1</td>
<td>Short-term</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>California</td>
</tr>
<tr>
<td>10</td>
<td>25.0%</td>
<td>Medium</td>
<td>Rural</td>
<td>1</td>
<td>1</td>
<td>Short-term</td>
<td>Day</td>
<td>Right</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Indiana</td>
</tr>
<tr>
<td>11</td>
<td>17.4%</td>
<td>Light</td>
<td>Rural</td>
<td>1</td>
<td>1</td>
<td>Short-term</td>
<td>Night</td>
<td>3.6</td>
<td>–</td>
<td>Short</td>
<td>Clear</td>
<td>–</td>
<td>Yes</td>
<td>–</td>
<td>–</td>
<td>South Carolina</td>
</tr>
<tr>
<td>12</td>
<td>–</td>
<td>Heavy</td>
<td>Urban</td>
<td>2</td>
<td>1</td>
<td>Short-term</td>
<td>Day</td>
<td>Left</td>
<td>–</td>
<td>Clear</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>San Antonio</td>
</tr>
<tr>
<td>13</td>
<td>4.5%</td>
<td>Medium</td>
<td>Rural</td>
<td>1</td>
<td>1</td>
<td>Long-term</td>
<td>Night</td>
<td>3.3</td>
<td>Right</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>North Carolina</td>
</tr>
<tr>
<td>14</td>
<td>–</td>
<td>–</td>
<td>Urban</td>
<td>2</td>
<td>1</td>
<td>Long-term</td>
<td>Day</td>
<td>3.6</td>
<td>Left</td>
<td>Right</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Toronto</td>
</tr>
<tr>
<td>15</td>
<td>5.0%</td>
<td>Heavy</td>
<td>Urban</td>
<td>2</td>
<td>1</td>
<td>Long-term</td>
<td>Day</td>
<td>3.6</td>
<td>Left</td>
<td>Left</td>
<td>Clear</td>
<td>–</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>Florida</td>
</tr>
<tr>
<td>16</td>
<td>5.0%</td>
<td>Heavy</td>
<td>Urban</td>
<td>2</td>
<td>1</td>
<td>Long-term</td>
<td>Day</td>
<td>3.6</td>
<td>Right</td>
<td>Clear</td>
<td>–</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Florida</td>
</tr>
<tr>
<td>17</td>
<td>–</td>
<td>Heavy</td>
<td>Rural</td>
<td>2</td>
<td>2</td>
<td>Long-term</td>
<td>Day</td>
<td>–</td>
<td>Left</td>
<td>Left</td>
<td>Clear</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Wisconsin</td>
</tr>
<tr>
<td>18</td>
<td>–</td>
<td>Heavy</td>
<td>Rural</td>
<td>1</td>
<td>2</td>
<td>Long-term</td>
<td>Day</td>
<td>–</td>
<td>Left</td>
<td>Clear</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Pearland</td>
</tr>
</tbody>
</table>

Note: The symbol “–” represents the “unknown” value assigned for the unavailable variable in a work zone dataset.
4.5 Model Evaluation

In order to evaluate the accuracy of the decision tree-based capacity estimation model, it is compared with the existing capacity estimation models and the HCM (2000) using 18 sets of evaluation data given in Table 4.1. Table 4.3 shows the detailed input values of the evaluation datasets. It should be noted that twelve work zone sites (#1-12) in Table 4.3 are short-term and the remaining six sites (#13-18) are long-term.

Krammes and Lopez (1994) proposed the following model to estimate work zone capacity:

\[
C = (1600 \text{pcphpl} + I - R) \times H
\]  

(4.6)

where \( C \) is the predicted work zone capacity (vphpl); \( I \) is the adjustment value of the work intensity, which is suggested that 160 for low intensity, 0 for medium intensity and -160 for heavy intensity (Karim and Adeli, 2003); \( R \) is the adjustment value of the presence of ramps (\( R=0 \) if no ramp is present and \( R=160 \) if entrance ramp is present); \( H \) is the heavy vehicle adjustment factor.

Kim et al. (2001) presented the following multiple regression model based on 12 sets of Maryland short-term work zone data:

\[
C = 1857 - 168.1 \text{NUMCL} - 37.0 \text{LOCCL} - 9.0HV + 92.7LD - 34.3WL - 106.1WI - 2.3WG \times HV
\]  

(4.7)

where \( \text{NUMCL} \) is the number of closed lanes, \( \text{LOCCL} \) is the lane closure location (right=1, left=0), \( HV \) is the heavy vehicle percentage, \( LD \) is the lateral distance, \( WL \) is the work zone length, \( WI \) is the work intensity, \( WG \) is the work zone grade.

For long-term work zones, Al-Kaisy and Hall (2003) developed a generic multiplicative model taking into account several adjustment factors. This model is
written as:

\[ C = C_h \times f_{HV} \times f_d \times f_w \times f_s \times f_r \times f_l \times f_i \]  \hspace{1cm} (4.8)

where \( C_h \) = based work zone capacity (2000 pcphpl), \( f_{HV} \) = adjustment factor for heavy vehicles, \( f_d \) = adjustment factor for driver population, \( f_w \) = adjustment factor for work activity, \( f_s \) = adjustment factor for side of lane closure, \( f_r \) = adjustment factor for rain, \( f_l \) = adjustment factor for light condition, \( f_i \) = adjustment factor for non-additive interactive effects.

Recently, Heaslip et al. (2009) developed the following analytical model to estimate the capacity of a long-term freeway work zone:

\[ C = f_l \times f_d \times f_r \times (C_{unadj} - V_r) \]  \hspace{1cm} (4.9)

where \( f_l \) = adjustment factor for the lighting condition, \( f_d \) = adjustment factor for the driver population, \( f_r \) = adjustment factor for rain, \( V_r \) = ramp volume and \( C_{unadj} \) is the unadjusted capacity (vphpl) that can be estimated by three equations proposed in their study.

Table 4.4 summarizes the work zone capacity estimation results from the decision tree-based model as well as the above four existing capacity estimation models. It should be noted that the two short-term work zone capacity models in Eqs. (4.6)-(4.7) are not used to estimate the capacities of long-term work zones because they may yield significant estimated errors for the long-term work zones. Similarly, the two long-term work zone capacity models in Eqs. (4.8)-(4.9) are not applicable for short-term work zones. In this chapter, the decision tree-based model has the capability of estimating short-term and long-term work zone capacities.
Table 4.4 Comparison of estimated work zone capacities from the decision tree-based model with four existing capacity estimation models

<table>
<thead>
<tr>
<th>State/City</th>
<th>Site #</th>
<th>Observed capacity (vphpl)</th>
<th>Estimated capacity (vphpl)</th>
<th>RMSE=</th>
<th>MPE=</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Krammes &amp; Lopez (1994)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maryland</td>
<td>1</td>
<td>1408</td>
<td>1297</td>
<td>1309</td>
<td>N.A</td>
</tr>
<tr>
<td>Texas</td>
<td>2</td>
<td>1641</td>
<td>1705</td>
<td>1612</td>
<td>N.A</td>
</tr>
<tr>
<td>Texas</td>
<td>3</td>
<td>1522</td>
<td>1555</td>
<td>1574</td>
<td>N.A</td>
</tr>
<tr>
<td>Texas</td>
<td>4</td>
<td>1402</td>
<td>1622</td>
<td>1443</td>
<td>N.A</td>
</tr>
<tr>
<td>Texas</td>
<td>5</td>
<td>1668</td>
<td>1697</td>
<td>1612</td>
<td>N.A</td>
</tr>
<tr>
<td>Texas</td>
<td>6</td>
<td>1180</td>
<td>1427</td>
<td>1337</td>
<td>N.A</td>
</tr>
<tr>
<td>Texas</td>
<td>7</td>
<td>1505</td>
<td>1571</td>
<td>1611</td>
<td>N.A</td>
</tr>
<tr>
<td>Texas</td>
<td>8</td>
<td>1490</td>
<td>1429</td>
<td>1505</td>
<td>N.A</td>
</tr>
<tr>
<td>California</td>
<td>9</td>
<td>1200</td>
<td>1286</td>
<td>1399</td>
<td>N.A</td>
</tr>
<tr>
<td>Indiana</td>
<td>10</td>
<td>1500</td>
<td>1280</td>
<td>1467</td>
<td>N.A</td>
</tr>
<tr>
<td>South Carolina</td>
<td>11</td>
<td>1440</td>
<td>1363</td>
<td>1610</td>
<td>N.A</td>
</tr>
<tr>
<td>San Antonio</td>
<td>12</td>
<td>1270</td>
<td>1286</td>
<td>1399</td>
<td>N.A</td>
</tr>
<tr>
<td>North Carolina</td>
<td>13</td>
<td>1787</td>
<td>N.A</td>
<td>1627</td>
<td>1798</td>
</tr>
<tr>
<td>Toronto</td>
<td>14</td>
<td>1644</td>
<td>N.A</td>
<td>1634</td>
<td>1455</td>
</tr>
<tr>
<td>Florida</td>
<td>15</td>
<td>2090</td>
<td>N.A</td>
<td>1738</td>
<td>1778</td>
</tr>
<tr>
<td>Florida</td>
<td>16</td>
<td>1972</td>
<td>N.A</td>
<td>1610</td>
<td>1778</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>17</td>
<td>1896</td>
<td>N.A</td>
<td>1610</td>
<td>1856</td>
</tr>
<tr>
<td>Pearland</td>
<td>18</td>
<td>1670</td>
<td>N.A</td>
<td>2000</td>
<td>1642</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Short-term (#1-12)</td>
<td>128</td>
<td>108</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Long-term (#13-18)</td>
<td>280</td>
<td>170</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+2.1%</td>
<td>-6.7%</td>
<td>-6.6%</td>
<td>+0.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+4.3%</td>
<td>-6.7%</td>
<td>-6.6%</td>
<td>+0.9%</td>
</tr>
</tbody>
</table>

According to the statistical comparison results in Table 4.4, it can be seen that the root mean square error (RMSE) of short-term work zone capacity from the decision tree-based model is 66, substantially lower than the RMSE values of 128 and 108 obtained from the models of Krammes and Lopez (1994) and Kim et al. (2001), respectively. In comparison to the existing two long-term work zone capacity models, the decision tree-based model also yields a far lower RMSE of 59 in estimating the long-term work zone capacity. In addition, the mean percent error (MPE) from the decision tree-based model for short-term work zones is -0.9%, compared with +2.1%.
from the model of Krammes and Lopez (1994), and +4.3% from the model of Kim et al. (2001). For the long-term work zones, the decision tree-based model yields a small positive $MPE$ (less than one percent), whereas Al-Kaisy et al. (2003) and Heaslip et al. (2009) models both yield a big $MPE$ (more than 6 percent). One possible reason for the higher $RMSE$ and $MPE$ generated by the existing models is that the collected work zone datasets are incomplete (e.g., the input values of some variables in these datasets are not given). Because of the incomplete inputs, the existing capacity models yield poor results. In summary, the statistical comparison results demonstrate that the decision tree-based model outperforms the existing short-term work zone capacity models in estimating short-term work zone capacity and performs better than the existing long-term work zone capacity models.

As mentioned earlier, the HCM provides the capacity estimation guidelines both for the short-term work zones and for the long-term work zones. For the purpose of estimation accuracy comparison, the HCM is also applied to estimate the capacities of work zones listed in Table 4.3. The estimates from the HCM are further compared with those from the decision tree-based model. Figure 4.4 (a) and (b) shows a comparison of the estimated work zone capacity from the decision tree-based model and the HCM. It can be seen that the overall $RMSE$ for all 18 evaluation work zone datasets from the HCM is 137, larger than double that from the decision tree-based model. The ideal lines in Figure 4.4 (a) and (b) represent a perfect correlation that estimated work zone capacities fully match the field data. In Figure 4.4 (b), the estimated results of long-term work zone from the HCM all are below the ideal line, suggesting that the HCM is more likely to underestimate long-term work zone capacity. As compared with the slight departure of the scattered dots shown in Figure 4.4 (a), the widely scattered capacity dots from the ideal line in Figure 4.4 (b)
indicates that the HCM performs much worse than the decision tree-based model. Using the testing data provided in the study of Adeli and Jiang (2003), it can be found that the decision tree-based model provides a RMSE of 105, less than the RMSE of 127 from the neural-fuzzy model addressed by Adeli and Jiang (2003). Therefore, it can be concluded that the developed decision tree-based model is a good alternative to estimate the short-term and long-term work zone capacity because it can estimate work zone capacity with a high degree of accuracy.

![Comparison of estimated and observed work zone capacity](image)

**Figure 4.4 Comparison of estimated and observed work zone capacity**

### 4.6 Summary

Numerous factors, such as heavy vehicle percentage, work time, weather and road type, are found to affect work zone capacity. Some of these influencing factors are categorical and some have impacts on work zone capacity by interacting with others. It is thus difficult to estimate capacity using a simple mathematical model. Decision tree is a useful and effective method to estimate work zone capacity because it partitions the space into many smaller regions and fits simple models to them,
which reduces estimated errors. Therefore, a decision tree-based capacity estimation model has been developed in this chapter to estimate work zone capacity. The $F$-test splitting criterion is employed to split node for tree growing. A post-pruning approach, which aims to minimize the prediction mean square error of checking data, is employed to prune the grown decision tree. The work zone capacity data collected from fourteen states and cities of U.S. are used to train, check and evaluate the decision tree. Finally, the decision tree comprising 16 leaf nodes has been selected as the optimal tree for estimating work zone capacity.

The statistical comparison results demonstrate that the decision tree-based model outperforms the existing two short-term and two long-term work zone capacity models, especially when the input values of variables are partially available for work zone capacity estimation. Unlike the HCM, the decision tree-based model does not require traffic engineers to set various adjustment factors based on their experience. Using this model, traffic engineers can easily estimate the capacity of a given work zone by referring to the straightforward “if-then” rules or by tracing a path down the tree to a leaf node. It avoids the estimated errors caused by their subjective judgment.

Therefore, the deterministic queuing model could provide an accurate estimate for traffic delay by using the work zone capacity derived from the decision tree-based model. As mentioned, work zone costs are usually high and there is a need to find an optimal subwork zone operational strategy for the minimization of total work zone costs. Hence, a total work zone cost minimization model will be developed to determine an optimal subwork zone strategy in next chapter. The deterministic queuing model incorporated in the total work zone cost minimization model will use the work zone capacity from the decision tree-based model.
CHAPTER 5  OPTIMAL SUBWORK ZONE OPERATIONAL STRATEGY FOR SHORT-TERM WORK ZONE PROJECTS

5.1 Introduction

Singapore is a “Garden City” with 1.3 million trees located in 300 parks and on more than 2,400 hectares of roadsides. Luxuriant trees are therefore everywhere along roadsides in Singapore. Figure 5.1 depicts two pictures shot from a highway in central business district (CBD) and a freeway from the Changi international airport to the city. These two pictures clearly show that luxuriant trees have actually formed a tree tunnel, and they should be trimmed regularly for the purpose of traffic safety.

![Orchard Road (CBD)](image1)

![East Cost Parkway (ECP) Freeway](image2)

Figure 5.1 Two roadway pictures shot in Singapore

To trim roadside trees in Singapore, one or two lanes adjacent to roadside have to be closed temporally and tree-trimming activities should be implemented in daytime due to local safety regulations. For a roadside tree-trimming project, the length of roadside trees to be trimmed is always known and the project is usually
required to be completed within a predetermined time window. According to the MUTCD (2003), the roadside tree-trimming project can be classified as a short-term work zone project, where daytime work occupies a location for more than one hour within a single daylight period. Hereafter, a short-term work zone project refers to a project with the similar characteristics of the roadside tree-trimming projects.

There are two possible operational strategies including: i) the single work zone operational strategy and ii) the subwork zone operational strategy to complete a short-term work zone project. A single work zone operational strategy closes the entire lane in one travel direction during the project implementation period. A subwork zone operational strategy, however, divides the entire work zone into multiple subwork zones, and the relevant work zone activities are undertaken on one subwork zone with partial lane closure at a time. Figure 5.2 schematically depicts these two work zone strategies. Figure 5.2(a) shows a single work zone operational strategy for a work zone project with the length of $L$ kilometers, and Figure 5.2(b) gives a five-subwork zone operational strategy.

As mentioned, work zone project costs including maintenance cost and user delay cost are usually high. Therefore, there is a need to find an optimal operational strategy to minimize work zone cost. Compared with a single work zone operational strategy, a subwork zone operational strategy is more realistic. This chapter therefore focuses on the total work zone cost minimization model development and algorithm design to find an optimal subwork zone operational strategy. Here, a subwork zone operational strategy comprises three elements: number of subwork zones, subwork
zone length and project start time. Since the HCA model in Chapter 3 could provide a high estimation accuracy of traffic delay, it will be incorporated in the total work zone cost minimization model. Because of its simplicity, the deterministic queuing model will be also incorporated in the total work zone cost minimization model and it will use the work zone capacity from the decision tree-based model in Chapter 4.

5.2 Notations, Assumptions & Problem Statement

The problem can be stated as follows: given a short-term work zone project with a total length of $L$ kilometers on a four-lane two-way expressway, find an optimal subwork zone operational strategy that can minimize the total work zone cost from the systemic perspective. A subwork zone operational strategy consists of the number of subwork zones, each subwork zone length and the start time of work activities on each subwork zone. Let $n$ denote the number of subwork zones in total, and $x_i$ and $y_i$ respectively denote the activity area length and the start time of the $i^{th}$ subwork zone, $i = 1, 2, \ldots, n$. For the sake of presentation, these subwork zone lengths and start times are grouped into two row vectors:
A feasible subwork zone operational strategy can be represented by these two row vectors \((x, y)\) fulfilling the following length and time window constraints:

\[
\sum_{i=1}^{n} x_i = L
\]

\[
l_{\text{min}} \leq x_i \leq l_{\text{max}}, i = 1, 2, \ldots, n
\]

\[
u_0 \leq y_j \leq v_1
\]

where \(l_{\text{min}}\) and \(l_{\text{max}}\) are the minimum and maximum lengths of a subwork zone restricted by traffic engineers, respectively; \(u_0\) and \(u_1\) are respectively the earliest project start time and the latest completion time.

To estimate the total work zone cost caused by a feasible subwork zone operational strategy, the following assumptions are made in this chapter:

(A1) The traffic flow approaching work zone in vehicles per hour (vph) at any time \(t \in \left[u_0 + j\delta, u_0 + (j + 1)\delta\right], j = 0, 1, 2, \ldots\), denoted by \(f(t)\), is approximated by a known hourly traffic flow at the beginning time of this time interval, \(f(u_0 + j\delta)\), where \(\delta\) is length of a uniform time interval in the traffic flow distribution. It is also assumed that the traffic flow approaching work zone is equally distributed over each lane.

(A2) The proportion of trucks approaching work zone is assumed to be a constant \(P_r\).

(A3) During each time interval \(\left[u_0 + j\delta, u_0 + (j + 1)\delta\right], j = 0, 1, 2, \ldots\), vehicles...
traveling at the average speed of \( V_m^{u_j+\delta} \) km/h pass through work zone, and at the average speed of \( V_a^{u_j+\delta} \) km/h approach work zone. \( V_m^{u_j+\delta} \) and \( V_a^{u_j+\delta} \) both vary with traffic flow.

**(A4)** During each time interval \([u_0 + j\delta, u_0 + (j+1)\delta]\), \( j = 0, 1, 2, \ldots \), the short-term work zone capacity in terms of vehicles per hour is \( C_w^{u_j+\delta} \), which can be estimated by using the decision tree-based model in Chapter 4.

According to HCM (2000), the four-lane two-way freeway capacity outside work zone is \( C_0^{u_j+\delta} \), where \( C_0^{u_j+\delta} = 2800 f_d f_w f_g f_{hv} \). Here, \( f_d \) is the traffic directional distribution factor, \( f_w \) is the adjustment factor for lane width, \( f_g \) is the adjustment factor for the grade, which is equal to 1.0 at 0% grade, \( f_{hv} \) is the heavy vehicle factor, where \( f_{hv} = [1 + P_T (E_T - 1)]^{-1} \), \( E_T \) is the passenger-car equivalent for the truck.

**(A5)** There are no signals, intersections, interchanges and access points in the work zone.

**(A6)** The user delay in terms of vehicle hours can be converted into user delay cost by an average cost \( C_{vh} \) dollars per vehicle-hour.

**(A7)** The time required to complete the work activities of a subwork zone is a function of the subwork zone activity area length \( x \), denoted by \( h(x) \).

**(A8)** In this chapter, traffic control cost is referred to as the work zone maintenance cost, denoted by \( g(x) \), which is a function of subwork zone length \( x \). It includes the rental cost of traffic control devices and expense for traffic control device relocation, installation, maintenance and removal.
After the work has been completed on one subwork zone, it is immediately moved to the next subwork zone. No work break is considered.

The lane closure configurations of all subwork zones are identical.

Note that the Assumption (A1) takes into account various traffic flow levels from low to high in a day. The Assumption (A3) applies varying traffic speeds in related to the traffic flow rate. It should be noted that the traffic speed is assumed as constant in previous studies. The Assumptions (A7) and (A8) use two general functions, \( h(x) \) and \( g(x) \), to express work time and traffic control cost for a subwork zone. The general work time function \( h(x) \) can be expressed by the following linear form:

\[
h(x) = d_1 + d_2 \times x
\]  

(5.6)

where \( d_1 \) is the fixed setup time for a subwork zone, \( d_2 \) is the variable work time per kilometer for a subwork zone.

The general function \( g(x) \) covers the total traffic control cost function used by McCoy and Mennenga (1998). It includes two parts: one is the cost for installing and removing traffic control devices and the other is the cost of traffic control devices themselves, which is expressed as a rental charge. Therefore, \( g(x) \) can be expressed by:

\[
g(x) = c_1 + c_2 \times x + (c_3 + c_4 \times x) \times h(x)
\]  

(5.7)

where \( c_1 \) is the fixed cost of installing and removing traffic control devices for altering and redirecting vehicles (S$); \( c_2 \) is the variable cost of installing and removing traffic control devices placed in the workspace (S$/km); \( c_3 + c_4 \times x \) is the cost for installing and removing traffic control devices; \( c_3 \) is the fixed cost of traffic
control devices for altering and redirecting vehicles (SS/h); \( c_4 \) is the variable cost of traffic control devices placed in the workspace (SS/km·h); \((c_3 + c_4 \times x) \times h(x)\) is the rental cost of traffic control devices.

5.3 Total Work Zone Cost Estimation

Given a feasible subwork zone operational strategy \((x,y)\), the total work zone cost, denoted by \( F_0(x,y) \), is the sum of costs of two different parties: contractors and road users. More specifically, it consists of three components: total user delay cost \( C_{DP}(x,y) \), total traffic control cost \( C_{TC}(x,y) \) and total additional traffic accident cost \( C_{AA}(x,y) \). Therefore, the total work zone cost can be calculated by:

\[
F_0(x,y) = C_{DP}(x,y) + C_{TC}(x,y) + C_{AA}(x,y)
\]  

The estimation of these three cost components is elaborated below.

5.3.1 Total User Delay Cost

This chapter assumes that the user delay cost is proportional to the user delay. Hence, the total user delay cost \( C_{DP}(x,y) \) is equal to the total user delay \( T_i(x,y) \) multiplied by the average delay cost per vehicle-hour \( C_{vh} \), namely,

\[
C_{DP}(x,y) = T_i(x,y) \times C_{vh}
\]

To estimate the total user delay \( T_i(x,y) \), the following deterministic queuing model (DQM) and the HCA model of Chapter 3 can be employed, respectively.
5.3.1.1 Deterministic queuing model (DQM)

For the subwork zone operational strategy \((x, y)\), the total user delay \(T(x, y)\) comprises the total queuing delay \(T_q(x, y)\) at upstream of subwork zones and the total moving delay \(T_m(x, y)\) when vehicles pass through subwork zones. That is, total user delay can be expressed by

\[
T(x, y) = T_q(x, y) + T_m(x, y)
\]

The detailed derivations of the total queuing delay \(T_q(x, y)\) and the total moving delay \(T_m(x, y)\) using the DQM are elaborated as follows.

5.3.1.1.1 Total queuing delay \(T_q(x, y)\) for the subwork zone operational strategy \((x, y)\)

For the \(i\)th subwork zone, the work time interval is \([y_i, y_i + h(x_i))\). Therefore, the number of time intervals covered by the \(i\)th subwork zone, denoted by \(N_i\), can be calculated by

\[
N_i = \left\lfloor \frac{y_i + h(x_i)}{\delta} \right\rfloor - \left\lfloor \frac{y_i}{\delta} \right\rfloor + 1, \quad i = 1, 2, \ldots, n
\]

where \(\lfloor z \rfloor\) is the maximum integer not exceeding that the positive number \(z\), \(\delta\) is a uniform time interval in the traffic flow distribution, \(y_i + h(x_i)\) is the completion time of the \(i\)th subwork zone. Since there is no work break, \(y_i + h(x_i)\) is also equal to the start time of the \((i+1)\)th subwork zone, \(y_{i+1}\).

In Figure 5.3 (a), the start time of \(k\)th time interval covered by the \(i\)th subwork zone, denoted by \(t_i(k)\), can be defined by
Chapter 5 Optimal Subwork Zone Operational Strategy for Short-term Work Zone Projects

\[
t_i(k) = \begin{cases} y_i, & \text{if } k = 0 \\ \left\lfloor \frac{y_i}{\delta} \right\rfloor + k \delta, & \text{if } k = 1, 2, \ldots, N_i - 1 \\ y_{i+1}, & \text{if } k = N_i \end{cases} \tag{5.12}
\]

Figure 5.3 Illustrations of three cases in the queuing delay estimation

For the \(i^{th}\) subwork zone, the DQM is now used to estimate the total queuing delay during the time interval \( [t_i(k), t_i(k+1)] \). The initial queue length for the \(k^{th}\) time interval, denoted by \(H_i(k)\), can be calculated by

\[
H_i(k) = \max \left\{ H_i(k-1) + \left[ f(t_i(k-1)) - C_w^{(k-1)} \right] \Delta t_i(k) - t_i(k-1), 0 \right\}, k = 1, \ldots, N_i \tag{5.13}
\]

where \(f(t_i(k-1))\) is the traffic flow rate during the \((k-1)^{th}\) time interval, \(C_w^{(k-1)}\) is the work zone capacity during the \((k-1)^{th}\) time interval.

Hence, the maximum queue length for the \(i^{th}\) subwork zone, denoted by \(H_i^{\max}(x, y)\), can be calculated by
\[ H_i^\text{max}(x,y) = \max \{ H_i(0), H_i(1), \ldots, H_i(N_i-1) \} \]

Because the accumulated queue in the current time interval may propagate to the next one, the following two cases should be considered:

**Case 1.** In the \( k \)th time interval, the traffic flow approaching work zone \( f(t_i(k)) \) exceeds the work zone capacity \( C_w^{(k)} \), namely:

\[ f(t_i(k)) > C_w^{(k)} \quad (5.15) \]

In this case, the accumulated queue will increase, as intuitively depicted by Figure 5.3 (b).

**Case 2.** The traffic flow approaching work zone does not exceed the work zone capacity, namely:

\[ f(t_i(k)) \leq C_w^{(k)} \quad (5.16) \]

In this case, the accumulated queue decreases or even disappears within the \( k \)th time interval, shown in Figure 5.3 (c). The duration of the queue lasting in this time interval can be calculated by

\[ \Delta t_i(k) = \frac{H_i(k)}{C_w^{(k)} - f(t_i(k))} \quad (5.17) \]

According to the DQM, the queuing delay during the \( k \)th time interval \([t_i(k), t_i(k+1)]\) (e.g., shadow area covered by the cumulative arrival and departure curves in Figure 5.3 (b) and (c)), denoted by \( Q_i(k) \), can be calculated by

\[ Q_i(k) = \begin{cases} 
\frac{H_i(k) + H_i(k+1)}{2} \times [t_i(k+1) - t_i(k)], & \text{for Case 1} \\
\frac{H_i(k)}{2} \times \Delta t_i(k), & \text{for Case 2}
\end{cases} \quad (5.18) \]

After obtaining the queuing delay at each time interval, the total queuing delay occurred in the \( i \)th subwork zone, denoted by \( t_q(x_i, y_i) \), can be estimated by
Chapter 5 Optimal Subwork Zone Operational Strategy for Short-term Work Zone Projects

\[ t_q(x_i, y_i) = \sum_{k=1}^{N-1} Q_i(k) \]  \hspace{1cm} (5.19)

The total queuing delay \( T_q(x, y) \) can thus be calculated by

\[ T_q(x, y) = \sum_{i=1}^n t_q(x_i, y_i) + \frac{1}{2} \left[ H_n \left( N_n \right) \right]^2 \]  \hspace{1cm} (5.20)

Note that the second term in the right hand side of Eq. (5.20) is the queuing delay of those vehicles that are in the queue at the completion time of last subwork zone.

5.3.1.1.2 Total moving delay \( T_m(x, y) \) for the subwork zone operational strategy \((x, y)\)

The total moving delay caused by the \( i^{th} \) subwork zone, denoted by \( t_m(x_i, y_i) \), is defined as the difference between travel times on the expressway with and without work zone. Hence, total moving delay in the \( k^{th} \) time interval, denoted by \( R_i(k) \), can be estimated by

\[ R_i(k) = \begin{cases} \left( \frac{x_i}{V_a^{(k)}} - \frac{x_i}{V_a^{(k)}} \right) \times C_w^{(k)} \times (t_i(k+1) - t_i(k)), & \text{for Case 1} \\ \left( \frac{x_i}{V_m^{(k)}} - \frac{x_i}{V_m^{(k)}} \right) \times f(t_i(k)) \times [t_i(k+1) - \Delta t_i(k)] + C_w^{(k)} \times \Delta t_i(k), & \text{for Case 2} \end{cases} \]  \hspace{1cm} (5.21)

where \( V_a^{(k)} \) is the approaching speed at upstream of work zone and \( V_m^{(k)} \) is the speed at work zone.

Therefore, the total moving delay \( T_m(x, y) \) for the subwork zone operational strategy \((x, y)\) can be expressed by

\[ T_m(x, y) = \sum_{i=1}^n \sum_{k=1}^{N_i} R_i(k) \]  \hspace{1cm} (5.22)

Note that the previous studies ignore a fact that a queue may completely
disappear before the end of a time interval when the arriving traffic flow rate at that time interval is lower than the work zone capacity. This could cause their queuing delay estimates at that time interval overestimated, compared to the correct calculation in Eq. (5.18) in this chapter. The moving delay estimation in Eq. (5.21) also remedies a flaw in existing moving delay estimation derived by Chen and Schonfeld (2006) and Tang and Chien (2008). This is because they assume that the vehicle departure rate is equal to the arriving traffic flow when the arriving traffic flow rate is lower than the work zone capacity. However, the vehicle departure rate may be greater than the arriving rate when there is a queue.

5.3.1.2 Heterogeneous Cellular Automata (HCA) model

As discussed in Chapter 3, the HCA model is also able to estimate traffic delay occurred in work zone. Hence, the HCA model is also applied to estimate the total user delay \( T_i(x,y) \) for the subwork zone operational strategy \( (x,y) \). According to Duffy and McAvoy (2009), there is no evidence to reject the null hypotheses that the means of acceleration and deceleration rates in work zones located in four-lane two-way road are the same as those in six-lane two-way road. Hence, it is reasonable to assume that the drive acceleration-deceleration behavior in the work zone located in a four-lane two-way expressway is similar to that in a six-lane two-way freeway. The randomization probability function for a work zone in the four-lane two-way freeway is the same as that for a six-lane two-way freeway work zone. The detailed procedures using the HCA model are shown as follows.
Step 1: Let \( i = 1 \) and initialize the HCA model by simulating the HCA model for about 5 minutes.

Step 2: To calculate the user delay for the \( i^{th} \) subwork zone, the parameters including the activity area length of \( x_i \), the transition area length of \( L_t \), the light vehicle flow \( f_L (= f(t(0)) \cdot (1 - P_t) \) ), the heavy vehicle flow \( f_H (= f(t(0)) \cdot P_t) \), the maximum speed outside and in work zone, and the other essential parameters in Table 3.2 are input to the HCA model. The HCA model will be simulated for about a time duration of \( h(x_i) \) and record the travel delay of each vehicle traveling across the \( i^{th} \) subwork zone during the time period \([y_i, y_i + h(x_i)]\).

Step 3: If \( i \leq n \), \( n = n + 1 \) and then go to Step 2; Otherwise, go to Step 4.

Step 4: When the \( n^{th} \) subwork zone is just completed, the travel delays of those vehicles still queuing at upstream of the \( n^{th} \) subwork zone are recorded.

Step 5: The sum of the travel delays of all vehicles recorded at Step 2 and Step 4 is regarded as the total user delay \( T_i(x, y) \).

### 5.3.2 Total Traffic Control Cost & Additional Traffic Accident Cost

According to Assumption (A8), the total traffic control cost \( C_{TC}(x, y) \) can be expressed by

\[
C_{TC}(x, y) = \sum_{i=1}^{n} g(x_i)
\]  
(5.23)

The total additional traffic accident cost considered in this chapter is the cost of traffic
accidents occurring in the work zone and queue areas in the freeway. The traffic delay and the costs of traffic accidents occurred in the routes adjacent to the freeway are not taken into account. Similar to Jiang and Adeli (2003), the total additional traffic accident cost $C_{ad}(x, y)$ is equal to the number of accidents, $n_a$, per 100 million vehicle hours multiplied by the total user delay, $T_l(x, y)$, and the average cost per accident, $c_a$, namely,

$$C_{ad}(x, y) = \frac{n_a \times c_a \times T_l(x, y)}{10^9}$$  (5.24)

### 5.4 Total Work Zone Cost Minimization Model & Its Variation

The optimal subwork zone strategy with the uniform subwork zone length and time window constraints for a short-term work zone project can be determined by solving the total work zone cost minimization model as follows:

$$\min F_0(x, y) = C_{dp}(x, y) + C_{tc}(x, y) + C_{ad}(x, y)$$  (5.25)

subject to

$$\sum_{i=1}^{n} x_i = L$$  (5.26)

$$x_i = \frac{L}{n}, i = 1, 2, \ldots, n$$  (5.27)

$$l_{\min} \leq x_i \leq l_{\max}, i = 1, 2, \ldots, n$$  (5.28)

$$y_i = y_{i-1} + h(x_{i-1}), i = 2, \ldots, n$$  (5.29)

$$u_0 \leq y_i \leq u_i, i = 1, 2, \ldots, n$$  (5.30)

$$y_n + h(x_n) \leq u_i$$  (5.31)
where \( l_{\text{min}} \) and \( l_{\text{max}} \) are the minimum and maximum lengths of a subwork zone restricted by traffic engineers, respectively. The uniform subwork zone length constraint, shown by Eq. (5.27), ensures that each subwork zone has the uniform length. Eq. (5.29) implies that the \( i^{th} \) subwork zone is carried out immediately after the completion time of \((i-1)^{th}\) subwork zone. Eqs. (5.30) and (5.31) make sure that the work zone project is completed within the given time window.

From a theoretical perspective, adopting unequal subwork zone length may cause lower total work zone cost. A variation of the total work zone cost minimization model without the uniform subwork zone length constraint is presented below:

\[
\min F_0 (x, y) = C_{DP} (x, y) + C_{TC} (x, y) + C_{ud} (x, y) \quad (5.32)
\]

subject to

\[
\sum_{i=1}^{n} x_i = L \quad (5.33)
\]

\[
l_{\text{min}} \leq x_i \leq l_{\text{max}}, i = 1, 2, \ldots, n \quad (5.34)
\]

\[
y_i = y_{i-1} + h(x_{i-1}), i = 2, \ldots, n \quad (5.35)
\]

\[
u_0 \leq y_i \leq u_i, i = 1, 2, \ldots, n \quad (5.36)
\]

\[
y_n + h(x_n) \leq u_i \quad (5.37)
\]

It can be seen that the total user delay cost in Eq. (5.9) is a piecewise function because of the time-dependent traffic flow pattern. Therefore, the objective functions of the above two minimization models are non-differentiable functions with respect to project start time and subwork zone length. Moreover, the number of subwork zones is an implicit integer decision variable. Hence, these two models are the mixed-integer
non-differentiable minimization problems.

5.5 Solution Algorithm

Since the above two models are the mixed-integer non-differentiable minimization problems, they cannot be solved by conventional exact solution algorithms such as branch-and-bound methods developed for mixed-integer programming problems (Wolsey, 1998). Fortunately, meta-heuristic methods including the simulated annealing (SA) method and the genetic algorithm (GA) can be employed to solve these models because they only need to evaluate the objective function value and constraints at each iteration.

GA is very efficient in searching a large region of the solution space because the GA maintains a population of solutions and its crossover operator causes a large jump in the solution space. However, one limitation of the GA is that it cannot search the local region efficiently of the solution space since it has no explicit ways to produce a sequence of small moves in the solution space. SA is a stochastic iterative improvement method and it attempts to avoid being trapped in the local optimum by sometimes moving in a locally worse direction. However, it also has a limitation that SA cannot cover a large region of the solution space within a limited computation time since it is based on small moves. According to Koakutsu et al. (1996), the genetic simulated annealing (GSA) method outperforms GA and SA because it incorporates the merits of GA and SA while it avoids the limitations of GA and SA. Therefore, this chapter employs the GSA to solve the developed minimization models. The GSA
adopted in this chapter includes two phases: GA and SA. The GA phase generates a set of new solutions using the crossover and mutation operators and then the SA phase further refines each solution in the population.

To apply the GSA for solving the minimization model (5.32)-(5.37), there is a need to encode a subwork zone operational strategy into a chromosome. According to the constraints (5.33)-(5.34), the number of subwork zones should satisfy the relation:

\[
\left\lfloor \frac{L}{L_{\text{max}}} \right\rfloor \leq n \leq \left\lceil \frac{L}{L_{\text{min}}} \right\rceil
\]  

(5.38)

In light of the constraints expressed by Eq. (5.35) and Eq.(5.38), a subwork zone operational strategy \((x,y)\) can be encoded by the chromosome:

\[
z_n = x_1 x_2 \cdots x_n y_1
\]  

(5.39)

After the mutation and crossover operations, the induced chromosome may yield an infeasible subwork zone operational strategy, namely, it violates at least one of the constraints shown by Eqs. (5.33)-(5.37). It is therefore necessary to have a procedure that can repair the current chromosome yielding an infeasible solution.

### 5.5.1 Procedure to Repair Chromosome

**Step 1:** [Check the subwork zone length constraint] Sum up all genes \(x_i\) and yield that 

\[
S = \sum_{i=1}^{n} x_i
\]

If \(S = L\), go to Step 2. Otherwise, perform the following manipulations on the string \(x_1, \cdots, x_n\).

**Case 1:** If \(S > L\), find the maximum value in the current genes, say \(i\), and then gene \(x_i\) is updated by
\[ \bar{x}_i = x_i - \min \left[ \left(x_i - l_{\text{min}}, (S - L) \right) \right] \] (5.40)

The chromosome thus becomes \( z_n = x_1 x_2 \cdots \bar{x}_i \cdots x_n y_1 \), and then go to Step 1.

**Case 2:** If \( S < L \), find the minimum value in current genes, say \( x_i \), gene \( x_i \) is updated by
\[ \bar{x}_i = x_i + \min \left[ (l_{\text{max}} - x_i), (L - S) \right] \] (5.41)

The chromosome becomes \( z_n = x_1 x_2 \cdots \bar{x}_i \cdots x_n y_1 \), and go to Step 1.

**Step 2:** [Check the time window constraints (5.36)-(5.37)] If \( y_1 > u_1 - \sum_{i=1}^{n} h(x_i) \),

randomly generate a real number \( \bar{y}_1 \) in time interval \( [u_0, u_1 - \sum_{i=1}^{n} h(x_i)] \).

The chromosome becomes \( z_n = x_1, \cdots, x_n \bar{y}_1 \) and output the new chromosome.

Step 1 aims to repair the chromosome that generates a subwork operational strategy violating the total work zone length constraint expressed by Eq.(5.33). Step 2 repairs the chromosome such that the project start time is within the time window. Using the above chromosome repairing procedure, the following iterative algorithm embedding with the GSA method for solving the minimization model (5.32)-(5.37) is developed.
5.5.2 An Iterative Algorithm Embedding With the Genetic Simulated Annealing (GSA) Method

**Step 0:** (Initialization) Let the number of subwork zones, \( n = \max \left\{ \left\lfloor L / l_{\text{max}} \right\rfloor, 1 \right\} \).

**Step 1:** (Initialize the population) Generate a population of \( M \) chromosomes, denoted by

\[
P_n = \left\{ z_n^{(m)} = (x_1^{(m)}, \ldots, x_n^{(m)}, y_1^{(m)}), m = 1, 2, \ldots, M \right\}
\] (5.42)

where numbers \( x_i^{(m)}, i = 1, 2, \ldots, n \) are randomly generated in the interval \( [l_{\text{min}}, l_{\text{max}}] \), and \( y_i^{(m)} \) is also randomly generated in the time window \( [u_0, u_1] \).

**Step 2:** (Invoke the chromosome repairing procedure) Repair each chromosome in the population \( P_n \) to yield a new population \( \bar{P}_n \) in which each chromosome can be decoded into a feasible subwork zone operational strategy.

**Step 3:** (Calculate the fitness value for each chromosome). Decode each chromosome \( z_n^{(m)} = (x_1^{(m)}, \ldots, x_n^{(m)}, y_1^{(m)}), m = 1, 2, \ldots, M \) in population \( \bar{P}_n \) into a subwork zone operational strategy according to the scheme:

\[
x^{(m)} = (x_1^{(m)}, \ldots, x_n^{(m)})
\] (5.43)

\[
y^{(m)} = (y_1^{(m)}, \ldots, y_n^{(m)})
\] (5.44)

where

\[
y_i^{(m)} = y_{i-1}^{(m)} + h(x_{i-1}^{(m)}), i = 2, \ldots, n
\] (5.45)

Let the fitness value of chromosome \( z_n^{(m)} \) be equal to the total work zone...
cost for the subwork zone operational strategy shown in Eqs. (5.43)-(5.44), namely,

\[ F(z^{(m)}) = F_0(x^{(m)}, y^{(m)}), m = 1, 2, \ldots, M \]  

(5.46)

**Step 4:** (Initialize local and global optimums) Let \( z^{*(Local)} \) be a local best chromosome in the population \( P_n \) in terms of the minimum fitness value, and let the global best chromosome \( z^{*(Global)} = z^{*(Local)} \). Set the number of generations \( j = 1 \).

**Step 5:** (Check GSA stopping criterion) If \( j \leq N \), where \( N \) is a predetermined number of generations, go to Step 9. Otherwise, go to Step 6.

**Step 6:** (Replace the worst chromosome) Randomly choose two parent chromosomes \( z^{(m_1)} \) and \( z^{(m_2)} \) from the current population \( P_n \) with \( F(z^{(m_1)}) \neq F(z^{(m_2)}) \) and then do a single point crossover operation to reproduce two child chromosomes. With a preset crossover probability, one gene randomly selected from one parent chromosome is swapped with the corresponding gene on the other parent chromosome. It should be noted that when the swap does not occur, these two parent chromosomes are directly transferred to the child chromosomes. The worst chromosome \( z^{(m_0)} \) in the current population is replaced by the best child chromosome reproduced after the crossover operation, namely,

\[ z^{(m_0)} = \text{Crossover}(z^{(m_1)}, z^{(m_2)}) \]  

(5.47)

**Step 7.1:** (Initialize the temperature used by SA) Let the initial temperature \( T = T_0 \), where \( T_0 \) is a given initial temperature.
Step 7.2: (Check the SA stopping criterion) If $T \leq T_f$, where $T_f$ is a given low temperature, go to Step 8. Otherwise, go to Step 7.2.1

Step 7.2.1: (SA-based local search) A min-max mutation operator is applied to mutate the chromosome $z_n^{(m)}$ to reproduce a new chromosome:

$$\tilde{z}_n^{(m)} = \text{Mutate} \left( z_n^{(m)} \right).$$

During the mutation, all genes have a mutation probability of being mutated. For a gene selected for mutation, the value of this gene will be randomly changed into its minimum or maximum value. After the chromosome mutation, the chromosome repairing procedure is applied for the mutated chromosome $\tilde{z}_n^{(m)}$ to generate a feasible chromosome $\tilde{z}_n^{(m)}$.

Step 7.2.2: (SA-based local search) Calculate the energy change

$$\Delta E = F(\tilde{z}_n^{(m)}) - F(z_n^{(m)}).$$

If $\Delta E < 0$, then let $\tilde{z}_n^{(m)} = z_n^{(m)}$. Otherwise, randomly generate a random number $r$ in the interval $[0, 1]$, if $r < \exp(-\Delta E / T)$, also let $\tilde{z}_n^{(m)} = z_n^{(m)}$.

Step 7.2.3: (Replace the local minimum) If $F(\tilde{z}_n^{(m)}) < F(z_n^{*(\text{Local})})$, then

$$z_n^{*(\text{Local})} = \tilde{z}_n^{(m)}.$$

Step 7.2.4: (Update the temperature) Let $T = \alpha T$ where the predetermined parameter $0 < \alpha < 1$ and go to Step 7.2.

Step 8: (Update the population and global optimum). Update the current population $\bar{P}_n$ by replacing the chromosome $z_n^{(m)}$ with the local best-so-far chromosome $z_n^{*(\text{Local})}$, and if $F(z_n^{*(\text{Local})}) < F(z_n^{*(\text{Global})})$, then

$$z_n^{*(\text{Global})} = z_n^{*(\text{Local})}.$$

Let $j = j + 1$ and go to Step 5.
Step 9: (Check the stopping criterion) If \( n > \lfloor \frac{L}{L_{\text{min}}} \rfloor + 1 \), stop and output \( z_n^{(\text{Global})} \); otherwise, \( n = n + 1 \) and go to Step 1.

As for the first minimization model (5.25)-(5.31), the above solution algorithm is available by using the chromosome encoding scheme:

\[
z_n = \left( \frac{L}{n} \right) y_i
\]  

(5.48)

5.6 Numerical Example

To assess the proposed models and GSA solution algorithm, this chapter employs one example that is created from a short-term roadside tree-trimming project on a four-lane two-way freeway in Singapore. Suppose that the freeway facility has a posted speed limit of 90 km/h. The lane width is 3.5 meter. There are no traffic signals, intersections, interchanges and access points in the selected freeway segment. The selected expressway segment has a 0% grade. The tree-trimming project causes the fast lane closure in the selected expressway segment. The entire length of roadside trees to be trimmed in the project is 0.6 kilometers (\( L = 0.6 \text{ km} \)).

The parameters for the iterative algorithm embedding with the GSA method are set as follows: population size \( M = 100 \); the number of iterations used in the GSA stopping criterion (Step 5) \( N = 100 \); the initial temperature used by the SA (Step 7.1) \( T_0 = 100^\circ C \); the temperature in the SA stop criterion (Step 7.2) \( T_f = 0.1^\circ C \); the temperature updating parameter (Step 7.2.4) \( \alpha = 0.99 \); the crossover probability is 0.8 and the mutation probability is 0.1.

Since the tree-trimming activity must be implemented at daytime, the project
implementation time window is usually $[6:00, 18:00]$, namely, the earliest project start time is $u_0 = 6$ and the latest project completion time $u_1 = 18$. It is further assumed that the tree-trimming (work) time $h(x)$ takes the linear function expressed by Eq. (5.6) and the traffic control cost $g(x)$ takes the form expressed by Eq. (5.7).

It should be noted that these functions are already used by McCoy and Mennenga (1998). To estimate the parameters $d_1$, $d_2$, $c_1$, $c_2$, $c_3$ and $c_4$ involved in these two functions, I interviewed one manager of Swee Bee Contractor Pte Ltd (www.sweebee.sg) which has implemented a number of roadside tree-trimming projects in 2008 and obtained the following information:

(a) Tree trimming work rate and the setup time for a tree-trimming work zone are 3~10 trees/hour and 0~1.0 hours/zone, respectively;

(b) There are about 5~7 trees to be trimmed or pruned every 100 meters;

(c) Three advanced warning signs and about 50 traffic cones should be used to alert and redirect vehicles in expressway.

According to the above survey results and traffic control cost function used by McCoy and Mennenga (1998), this chapter takes a set of values for the parameters $d_1$, $d_2$, $c_1$, $c_2$, $c_3$ and $c_4$, as shown in Table 5.1. In addition, Table 5.1 gives the freeway capacity, the work zone capacity, the number of accidents per 100 million vehicle hours ($n_a$), average cost per accident ($c_a$) and average delay cost per vehicle-hour ($c_{dh}$).
### Table 5.1 Model parameters used in the numerical example

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Input values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>Fixed cost of installing and removing devices for alerting/redirecting vehicles</td>
<td>S$45</td>
</tr>
<tr>
<td>$c_2$</td>
<td>Variable cost of installing and removing traffic control devices in the workspace</td>
<td>S$70/km</td>
</tr>
<tr>
<td>$c_3$</td>
<td>Fixed cost of traffic control devices for altering/redirecting vehicles</td>
<td>S$30/h</td>
</tr>
<tr>
<td>$c_4$</td>
<td>Variable cost of traffic control devices in the workspace</td>
<td>S$45/km · h</td>
</tr>
<tr>
<td>$c_{a}$</td>
<td>Average cost per accident</td>
<td>S$142,000/acc</td>
</tr>
<tr>
<td>$c_{vh}$</td>
<td>Average delay cost per vehicle-hour</td>
<td>S$12/veh-hour</td>
</tr>
<tr>
<td>$d_1$</td>
<td>Fixed setup time for a subwork zone</td>
<td>0.2h</td>
</tr>
<tr>
<td>$d_2$</td>
<td>Variable work time per kilometer for a subwork zone</td>
<td>6h/km</td>
</tr>
<tr>
<td>$f_g$</td>
<td>Adjustment factor for grade</td>
<td>1.0</td>
</tr>
<tr>
<td>$f_w$</td>
<td>Adjustment factor for lane width</td>
<td>1.0</td>
</tr>
<tr>
<td>$f_d$</td>
<td>The traffic directional distribution factor</td>
<td>1.0</td>
</tr>
<tr>
<td>$l_{min}$</td>
<td>The minimum length of a subwork zone restricted by traffic engineers</td>
<td>0.1km</td>
</tr>
<tr>
<td>$l_{max}$</td>
<td>The maximum length of a subwork zone</td>
<td>0.6km</td>
</tr>
<tr>
<td>$n_a$</td>
<td>The number of accidents per 100 million vehicle hours</td>
<td>40acc/100mvh</td>
</tr>
<tr>
<td>$N_o$</td>
<td>The number of lanes open at work zone</td>
<td>1</td>
</tr>
<tr>
<td>$P_T$</td>
<td>Proportion of trucks</td>
<td>10%</td>
</tr>
<tr>
<td>$E_T$</td>
<td>Passenger-car equivalent for the truck</td>
<td>2.0</td>
</tr>
<tr>
<td>$u_0$</td>
<td>The earliest project start time</td>
<td>6</td>
</tr>
<tr>
<td>$u_1$</td>
<td>The latest completion time</td>
<td>18</td>
</tr>
<tr>
<td>$v_{a_{u_0+j}}$</td>
<td>Average speed of vehicles approaching work zone during time interval $[u_0+j, u_0+j+1]$</td>
<td>45km/h</td>
</tr>
<tr>
<td>$\delta$</td>
<td>The length of a uniform time interval in traffic flow distribution</td>
<td>1-hour</td>
</tr>
<tr>
<td>$C_{0_{u_0+j}}$</td>
<td>Freeway capacity during time interval $[u_0+j, u_0+j+1]$</td>
<td>2550vph</td>
</tr>
<tr>
<td>$C_{w_{u_0+j}}$</td>
<td>Work zone capacity during time interval $[u_0+j, u_0+j+1]$</td>
<td>1383vph*</td>
</tr>
</tbody>
</table>

* work zone capacity estimated from the decision tree-based model in Chapter 4

Using the HCM (2000) tabulated values and performing linear interpolation, the relationship between the traffic speed and the traffic flow can be expressed as:
\[
\text{Speed} = \begin{cases} 
45 \text{ (km/h)}, & \text{if } f/c > 1.0 \\
116.25 - 35.455 \times f/c \text{ (km/h)}, & \text{if } 0.89 < f/c \leq 1.0 \\
106.31 - 24.286 \times f/c \text{ (km/h)}, & \text{if } 0.68 < f/c \leq 0.89 \\
90.45 - 0.952 \times f/c \text{ (km/h)}, & \text{if } 0.47 < f/c \leq 0.68 \\
90 \text{ (km/h)}, & \text{if } f/c \leq 0.47 
\end{cases}
\] (5.49)

where \( f \) is the adjusted traffic flow rate and \( c \) is the capacity.

Assuming that the time interval \( \delta = 1 \) hour, Table 5.2 lists three time-dependent anticipated traffic flow patterns (traffic scenarios) and the time-dependent average speed of vehicles approaching work zone estimated according to Eq. (5.49). Note that the medium traffic scenario is the baseline scenario used for this example. The average speed of vehicles passing through work zone is also estimated, shown in Table 5.1.

**Table 5.2 Time dependent traffic flow and average travel speed**

<table>
<thead>
<tr>
<th>Time interval ( [u_0+j_u_0+j+1] )</th>
<th>Light traffic scenario</th>
<th>Medium traffic scenario*</th>
<th>Heavy traffic scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( f(u_0+j) ) ( V_{a+u}^{u+j} )</td>
<td>( f(u_0+j) ) ( V_{a+u}^{u+j} )</td>
<td>( f(u_0+j) ) ( V_{a+u}^{u+j} )</td>
</tr>
<tr>
<td></td>
<td>(vph) \ (km/h)</td>
<td>(vph) \ (km/h)</td>
<td>(vph) \ (km/h)</td>
</tr>
<tr>
<td>[6,7)</td>
<td>855 \ 90.0</td>
<td>1328 \ 90.0</td>
<td>1672 \ 89.8</td>
</tr>
<tr>
<td>[7,8)</td>
<td>1532 \ 89.8</td>
<td>2344 \ 86.2</td>
<td>2992 \ 78.0</td>
</tr>
<tr>
<td>[8,9)</td>
<td>765 \ 90.0</td>
<td>1193 \ 90.0</td>
<td>1496 \ 89.8</td>
</tr>
<tr>
<td>[9,10)</td>
<td>771 \ 90.0</td>
<td>1203 \ 90.0</td>
<td>1509 \ 89.8</td>
</tr>
<tr>
<td>[10,11)</td>
<td>711 \ 90.0</td>
<td>1113 \ 90.0</td>
<td>1392 \ 90.0</td>
</tr>
<tr>
<td>[11,12)</td>
<td>698 \ 90.0</td>
<td>1093 \ 90.0</td>
<td>1366 \ 90.0</td>
</tr>
<tr>
<td>[12,13)</td>
<td>971 \ 90.0</td>
<td>1503 \ 90.0</td>
<td>1899 \ 89.8</td>
</tr>
<tr>
<td>[13,14)</td>
<td>1171 \ 90.0</td>
<td>1803 \ 89.8</td>
<td>2289 \ 86.6</td>
</tr>
<tr>
<td>[14,15)</td>
<td>1181 \ 90.0</td>
<td>1818 \ 89.8</td>
<td>2309 \ 86.5</td>
</tr>
<tr>
<td>[15,16)</td>
<td>1527 \ 89.8</td>
<td>2337 \ 86.2</td>
<td>2983 \ 78.4</td>
</tr>
<tr>
<td>[16,17)</td>
<td>1638 \ 89.8</td>
<td>2503 \ 83.2</td>
<td>3199 \ 73.4</td>
</tr>
<tr>
<td>[17,18)</td>
<td>1438 \ 89.8</td>
<td>2204 \ 87.4</td>
<td>2810 \ 80.4</td>
</tr>
<tr>
<td>AHT</td>
<td>1101</td>
<td>1703</td>
<td>2163</td>
</tr>
<tr>
<td>Volume/capa</td>
<td>0.40</td>
<td>0.65</td>
<td>0.90</td>
</tr>
</tbody>
</table>

* Baseline traffic scenario in the numerical example.
5.6.1 Results

Since the number of iterations is equal to the number of subwork zones, the iterative algorithm embedding with the GSA needs six iterations for this example. For each iteration, the GSA method is applied to find the optimal subwork zone operational strategy with respect to a given number of subwork zones. From a practical perspective, work zone contractors always take the uniform length for each subwork zone (uniform length constraint). The first part of Table 5.3 presents the iterative results from the minimization model (5.25)-(5.31) which employs the DQM to estimate total user delay. It can be clearly seen that adopting two-subwork zones with the uniform length of 0.300 km and the project start time of 8:00 am is the optimal subwork zone operational strategy, which generates the minimum total work zone cost of S$492.6. The second part of Table 5.3 presents the optimal subwork zone operational strategy from the minimization model (5.32)-(5.37) without the uniform length constraint. The strategy that adopting two-subwork zones with the respective length of 0.292 km and 0.308 km and the project start time of 8:00 am is found to be the optimal subwork zone operational strategy and the resulting minimum total work zone cost is S$492.1. Compared with the optimal subwork zone operational strategy from the minimization model (5.25)-(5.31) subject to uniform length constraint, the optimal subwork zone operational strategy without the uniform length constraint has a lower total work zone cost though the total work zone cost decrement is marginal due to the low value of $c_{vb}$ for this particular example. This is because the optimal solution of the minimization model (5.25)-(5.31) is mathematically a feasible solution of the minimization model...
(5.32)-(5.37).

Table 5.3 Optimal subwork zone operational strategy with a given number of subwork zones at medium traffic scenario

<table>
<thead>
<tr>
<th>Number of subwork zones</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>0.600</td>
<td><strong>0.300</strong></td>
<td>0.200</td>
<td>0.150</td>
<td>0.120</td>
<td>0.100</td>
</tr>
<tr>
<td>$x_2$</td>
<td><strong>0.300</strong></td>
<td>0.200</td>
<td>0.150</td>
<td>0.120</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>$x_3$</td>
<td></td>
<td>0.200</td>
<td>0.150</td>
<td>0.120</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>$x_4$</td>
<td></td>
<td></td>
<td>0.150</td>
<td>0.120</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>$x_5$</td>
<td></td>
<td></td>
<td></td>
<td>0.120</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>$x_6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>Total work zone cost(SS): $F_0(x,y)$</td>
<td>657.7</td>
<td><strong>492.6</strong></td>
<td>503.3</td>
<td>610.5</td>
<td>792.9</td>
<td>1043.6</td>
</tr>
<tr>
<td>Project start time: $y_1$</td>
<td>8:10am</td>
<td><strong>8:00am</strong></td>
<td>8:00am</td>
<td>8:00am</td>
<td>8:00am</td>
<td>8:00am</td>
</tr>
</tbody>
</table>

Minimization model (5.32)-(5.37)*

<table>
<thead>
<tr>
<th>Number of subwork zones</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>0.600</td>
<td><strong>0.292</strong></td>
<td>0.199</td>
<td>0.153</td>
<td>0.120</td>
<td>0.100</td>
</tr>
<tr>
<td>$x_2$</td>
<td><strong>0.308</strong></td>
<td>0.199</td>
<td>0.138</td>
<td>0.120</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>$x_3$</td>
<td></td>
<td>0.202</td>
<td>0.157</td>
<td>0.120</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>$x_4$</td>
<td></td>
<td></td>
<td>0.152</td>
<td>0.120</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>$x_5$</td>
<td></td>
<td></td>
<td></td>
<td>0.120</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>$x_6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>Total work zone cost(SS): $F_0(x,y)$</td>
<td>657.7</td>
<td><strong>492.1</strong></td>
<td>502.9</td>
<td>608.5</td>
<td>792.9</td>
<td>1043.6</td>
</tr>
<tr>
<td>Project start time: $y_1$</td>
<td>8:10am</td>
<td><strong>8:00am</strong></td>
<td>8:00am</td>
<td>8:00am</td>
<td>8:00am</td>
<td>8:00am</td>
</tr>
</tbody>
</table>

* Total user delay is estimated by using the deterministic queuing model (DQM)

For the purpose of comparison, the HCA model is also used to determine the optimal subwork zone operational strategy in the minimization model (5.25)-(5.31).
The maximum speeds outside work zone and in work zone are respectively 50 cell/sec (=90km/h) and 25 cell/sec (=45km/h) for the HCA model. The transition area length is taken as 30 m. Table 5.4 presents the comparison results of optimal subwork zone operational strategy from the minimization model (5.25)-(5.31) using two different user delay estimation approaches. It can be seen that adopting two-subwork zones with the uniform length of 0.300 km and the project start time of 8:00 am is still the optimal subwork zone operational strategy if the HCA model is applied to estimate the total user delay. However, the corresponding minimum total work zone cost is S$524.2, larger than S$492.6 when the DQM is used.

| Estimation approach | $C_{TC}(x,y)$ | $C_{DP}(x,y)$ | $C_{AD}(x,y)$ | $F_{o}(x,y)$ | $x_i$ | $y_i$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DQM 1</td>
<td>303.6 S$</td>
<td>349.2 S$</td>
<td>4.9 S$</td>
<td>657.7 S$</td>
<td>0.60km</td>
<td>8:10am</td>
</tr>
<tr>
<td>HCA 1</td>
<td>303.6 S$</td>
<td>421.9 S$</td>
<td>5.9 S$</td>
<td>731.4 S$</td>
<td>0.60km</td>
<td>8:10am</td>
</tr>
<tr>
<td>DQM 2</td>
<td>306.0 S$</td>
<td>184.0 S$</td>
<td>2.6 S$</td>
<td>492.6 S$</td>
<td>0.30km</td>
<td>8:00am</td>
</tr>
<tr>
<td>HCA 2</td>
<td>306.0 S$</td>
<td>215.2 S$</td>
<td>3.0 S$</td>
<td>524.2 S$</td>
<td>0.30km</td>
<td>8:00am</td>
</tr>
<tr>
<td>DQM 3</td>
<td>340.8 S$</td>
<td>160.3 S$</td>
<td>2.2 S$</td>
<td>503.3 S$</td>
<td>0.20km</td>
<td>8:00am</td>
</tr>
<tr>
<td>HCA 3</td>
<td>340.8 S$</td>
<td>212.6 S$</td>
<td>3.0 S$</td>
<td>556.4 S$</td>
<td>0.20km</td>
<td>8:00am</td>
</tr>
<tr>
<td>DQM 4</td>
<td>383.7 S$</td>
<td>223.7 S$</td>
<td>3.1 S$</td>
<td>610.5 S$</td>
<td>0.15km</td>
<td>8:00am</td>
</tr>
<tr>
<td>HCA 4</td>
<td>383.7 S$</td>
<td>282.5 S$</td>
<td>4.0 S$</td>
<td>670.2 S$</td>
<td>0.15km</td>
<td>8:00am</td>
</tr>
</tbody>
</table>

From the table, it can be found that the total user delay from the HCA is a little larger than that from the DQM for any given number of subwork zones. The
underestimation of user delay from the DQM may be due to the neglect of the delays caused by the deceleration actions when vehicles approach work zone. Another possible reason could be that the DQM does not take into account the queuing delay when the approaching traffic flow is less than the work zone capacity. However, in reality, vehicle queues may form when traffic flow is below the work zone capacity because of the randomness of traffic flow.

According to Eq. (5.8), total work zone cost can be divided into three components: total user delay cost, total traffic control cost and total additional accident cost. These three costs vary with the number of subwork zones and their relationships are depicted in Figure 5.4. According to the figure, it can be seen that the total traffic control cost increases with the number of subwork zones. In other words, work zone contractors prefer to adopt the single work zone operational strategy because it causes the least traffic control cost. Figure 5.4 also shows that the total user delay cost curve likes a parabola. The minimum total user delay cost occurs at the three-subwork zone operational strategy. Similar to the total user delay cost, the total traffic accident cost also has a parabolic curve though it is close to the x-axis because their values are very small, ranging from 2S$ to 8S$. The relatively small value of traffic accident cost shows that it has a marginal effect on the optimal subwork zone strategy.
Figure 5.4 Three components of total work zone cost incurred by the optimal subwork zone operational strategy with a given number of subwork zones

From the road user’s viewpoint, the optimal three-subwork zone operational strategy is the best choice because it causes the least traffic delay cost. In reality, the optimal two-subwork zone operational strategy with the minimum total work zone cost, as shown in Table 5.3, is a compromise solution between work zone contractors and road users.

5.6.2 Impacts of $d_1$ & $d_2$

To assess the impact of the fixed work zone setup time $d_1$ and the variable work time rate $d_2$ on the number of subwork zones in the optimal subwork zone operational strategy formulated by the minimization model (5.25) - (5.31), the parameters $d_1$ and $d_2$ are varied. Note that the traffic control cost for a subwork zone, shown in Eq. (5.7), is also depending on these two parameters. Given a pair of
parameters $d_1$ and $d_2$, the proposed GSA solution algorithm can generate the minimum total work zone cost with respect to a given number of subwork zones, denoted by $F^*_0(d_1,d_2,n)$, where $n$ is the number of subwork zones. This chapter thus defines the following total work zone cost reduction ratio, $\rho(d_1,d_2,n)$, at a given number of subwork zones compared to the minimum total work zone cost caused by the single work zone operational strategy:

$$\rho(d_1,d_2,n) = \frac{F^*_0(d_1,d_2,n) - F^*_0(d_1,d_2,1)}{F^*_0(d_1,d_2,1)} \times 100\%$$

(5.50)

Figure 5.5 Impacts of parameters $d_1$ and $d_2$ on the optimal subwork zone operational strategy

Figure 5.5 illustrates the total work zone cost reduction ratio defined by Eq. (5.50) for various combinations of the parameters $d_1$ and $d_2$. Note that the shading column in the figure represents the optimal subwork zone operational strategy with respect to a given pair of $d_1$ and $d_2$. It can be seen that the number of subwork zones in the optimal subwork zone operational strategy gradually declines with the increase of the fixed work zone set up time $d_1$ and/or the variable work time rate $d_2$. 

131
Chapter 5 Optimal Subwork Zone Operational Strategy for Short-term Work Zone Projects

This is because a larger value for one of these two parameters results in a longer completion time of the project so that a part of tree trimming work has to be implemented under the heavy traffic conditions. The maximum total work zone cost reduction ratio is up to 29% when $d_1 = 0.1 \ h/\text{zone}$ and $d_2 = 5 \ h/\text{km}$. However, when $d_1 \geq 0.5 \ h/\text{zone}$ and $d_2 = 5 \ h/\text{km}$, or $d_1 = 0.1 \ h/\text{zone}$ and $d_2 \geq 8 \ h/\text{km}$, the total work zone cost reduction ratio is zero, namely, the single work zone operational strategy is the best solution. These results further confirm that the subwork zone operational strategy is more preferable as long as a subwork zone can be quickly set up and the work time rate is high.

5.6.3 Impact of Time Dependent Traffic Flow

To evaluate the impact of time dependent traffic flow on the optimal subwork zone operational strategy, this chapter also considers another two traffic scenarios in Table 5.2: light and heavy. The average hourly traffic flow (AHT) from 6:00 to 18:00 is 1101 vph for the light traffic scenario and 2163 vph for the heavy traffic scenario. The AHT and the corresponding average speed of vehicles approaching work zone for the two traffic scenarios are presented in Table 5.2. The volume/capacity ratios in three traffic scenarios are 0.40 for the light traffic scenario, 0.65 for the medium traffic scenario and 0.90 for the heavy scenario, respectively.
Figure 5.6 gives the number of sub-work zones in the optimal subwork zone operational strategy for each of the three traffic scenarios with respect to different fixed work zone setup time. According to the figure, it can be observed that the single work zone operational strategy is the best choice in the heavy traffic scenario (AHT=2163 vph, volume/capacity=0.90) if the fixed work zone set up time is not less than 0.1 h/zone. This is because the induced total user delay cost reduction cannot compensate the total traffic control cost increment as the number of subwork zones increases. This phenomenon is illustrated by Figure 5.7.
Figure 5.7 Three components of the total work zone cost with respect to a given number of subwork zones for the heavy traffic scenario when $d_1=0.1h/zone$.

Figure 5.6 also shows that the number of subwork zones in the optimal subwork zone operational strategy gradually declines for the medium traffic scenario (AHT=1703vph, volume/capacity=0.65) when the fixed work zone set up time increases. For example, when the fixed set up time is greater than or equal to 0.5 h/zone, the single work zone operational strategy is the optimal solution for the medium traffic scenario (in Figure 5.8). According to Figure 5.6 and 5.9, it can be also seen that the subwork zone operational strategy including two-subwork zones is the optimal strategy in the light traffic scenario (AHT=1101vph, volume/capacity=0.40), even if the fixed set up time is a big value.
Figure 5.8 Three components of the total work zone cost with respect to a given number of subwork zones for the medium traffic scenario when $d_1 = 0.5\, \text{h/zone}$

Figure 5.9 Three components of the total work zone cost with respect to a given number of subwork zones for the light traffic scenario when $d_1 = 0.5\, \text{h/zone}$
5.7 Summary

This chapter has proposed a more realistic subwork zone operational strategy issue by taking into account the variable traffic speed, time window and subwork zone length constraints. It has also remedied a few flaws existed in the previous queuing delay and moving delay estimation formulae. Subject to the time window and uniform length constraints, a non-differentiable minimization model has been built to minimize the total work zone cost. To investigate the impact of unequal subwork zone length constraint on the optimal subwork zone strategy, this chapter also presented a variation of the minimization model without the uniform length constraint. The iterative algorithm embedding with the genetic simulation annealing (GSA) method has been designed to solve the proposed two models, which includes an interesting chromosome-repairing procedure. The numerical example has shown the availability and flexibility of the proposed models in determining the optimal subwork zone operational strategy.

In reality, different parties have different objectives. For example, work zone contractors are more concerned with the total maintenance cost. However, drivers and land transportation authorities aim to minimize total user delay caused by a subwork zone operational strategy. In next chapter, the optimal subwork zone operational strategy by minimizing the total maintenance cost subject to the threshold constraint imposed on total user delay will be investigated.
6.1 Introduction

The total work zone cost in Chapter 6 can be regarded as the system cost that integrates costs of two different parties: work zone contractor and road user. However, different parties have different objectives. With the objective of profit maximization, work zone contractor cares about the minimization of total work zone maintenance cost while road users only concern their travel delay in work zone. These two objectives contradict each other to some extent. To concern the benefits of road users, land transport authorities impose two constraints on a work zone project: (1) average travel delay in work zone should not be greater than a limited threshold; (2) maximum queue length per lane in work zone cannot exceed a desirable threshold. For example, the Maryland State Highway Administration (MSHA, 2006) requires that the queue length and average travel delay caused by the work zone cannot exceed 1.0 mile/lane and 15 minute/vehicle, respectively. If the resulting queue length and average travel delay cannot meet these two thresholds, the contractor has to rent other lanes or pay for the penalties imposed by the land transport authorities.

Although there are considerable studies addressed for the subwork zone operational strategy problem, their objectives are to find an optimal operational strategy in order to minimize the total work zone cost from the systemic viewpoint.
Little effort has been made to help contractors find an optimal operational strategy that minimizes the total maintenance cost. From the contractor’s standpoint, this chapter proposes a total maintenance cost minimization model to determine the optimal subwork zone operational strategy subject to the queue length and travel delay constraints imposed by the land transport authorities.

### 6.2 Definitions & Assumptions

Let $n$ denote the number of subwork zones in total, and $x_i$ and $y_i$ respectively denote the length and start time of the $i^{th}$ subwork zone, $i = 1, 2, \cdots, n$. For the sake of presentation, these subwork zone lengths and start times are grouped into two row vectors:

$$\mathbf{x} = (x_1, x_2, \cdots, x_n) \quad (6.1)$$

$$\mathbf{y} = (y_1, y_2, \cdots, y_n) \quad (6.2)$$

where $y_1$ is the start time of the 1$^{st}$ subwork zone (namely, the project start time).

A feasible subwork zone operational strategy can be represented by these two row vectors $(\mathbf{x}, \mathbf{y})$ fulfilling the following length and time window constraints:

$$x_i = l = L / n, i = 1, 2, \cdots, n \quad (6.3)$$

$$l_{\text{min}} \leq x_i \leq l_{\text{max}}, i = 1, 2, \cdots, n \quad (6.4)$$

$$u_0 \leq y_i \leq u_1 \quad (6.5)$$

where $l_{\text{min}}$ and $l_{\text{max}}$ are the minimum and maximum lengths of a subwork zone restricted by traffic engineers, respectively; $l$ is the uniform subwork zone length; $u_0$ and $u_1$ are respectively the earliest project start time and the latest completion
time.

To estimate the optimal subwork zone operational strategy by minimizing the total maintenance cost, this chapter employs the ten assumptions given in Section 5.2. Another two assumptions (A11 and A12) are also made for the travel delay and queue length constraints, shown as follows:

(A11). Average travel delay in work zone should not be greater than a limited threshold.

(A12) The maximum queue length per lane in work zone cannot exceed a desirable threshold.

Let $T_a(x,y)$ and $Q_{\text{max}}(x,y)$ denote the average user travel delay and maximum queue length per lane caused by the operational strategy $(x,y)$, respectively. Their thresholds imposed by land transport authorities are denoted by $T_a$ minutes per vehicle (min/veh) and $Q_{\text{max}}$ vehicles per lane (veh/lane), respectively. Hence, a feasible subwork zone operational strategy should fulfill the following two constraints:

$$Q_{\text{max}}(x,y) \leq Q_{\text{max}} \quad (6.6)$$
$$T_a(x,y) \leq T_a \quad (6.7)$$

Note that $T_a(x,y)$ and $Q_{\text{max}}(x,y)$ can be estimated by using the DQM or the HCA model. For simplicity, this chapter uses the DQM in Chapter 5 to estimate the average travel delay $T_a(x,y)$ and the maximum queue length per lane $Q_{\text{max}}(x,y)$. The detailed derivations for $T_a(x,y)$ and $Q_{\text{max}}(x,y)$ are shown as follows.
The average user delay, $T_a(x,y)$, can be expressed by

$$T_a(x,y) = \frac{T_q(x,y) + T_m(x,y)}{\sum_{i=1}^{n} \sum_{k=1}^{N_i} f(t_i(k))(t_i(k+1) - t_i(k))} \quad (6.8)$$

The maximum queue length per lane, $Q_{max}(x,y)$, can be calculated by

$$Q_{max}(x,y) = \max \left\{ H^*_1(x,y), H^*_2(x,y), \ldots, H^*_n(x,y) \right\} \quad (6.9)$$

where $N_o$ is the number of open lanes in the work zone.

### 6.3 Total Maintenance Cost Minimization Model

For a specific subwork zone, the maintenance cost in this chapter takes the form expressed by Eq. (5.7). Therefore, the total maintenance cost caused by the subwork zone operational strategy $(x,y)$ can be expressed by

$$C_{MC}(x,y) = \sum_{i=1}^{n} g(x_i) \quad (6.10)$$

The optimal subwork zone operational strategy from the contractor’s standpoint can be determined by solving the following total maintenance cost minimization model:

$$\min C_{MC}(x,y) \quad (6.11)$$

subject to

$$x_i = L/n, \quad i = 1,2,\cdots,n \quad (6.12)$$

$$l_{min} \leq x_i \leq l_{max}, i = 1,2,\cdots,n \quad (6.13)$$

$$u_0 \leq y_i \leq u_i, i = 1,2,\cdots,n \quad (6.14)$$

$$Q_{max}(x,y) \leq Q_{max} \quad (6.15)$$

$$T_a(x,y) \leq T_a \quad (6.16)$$
Eq. (6.14) defines the time window constraint, which restricts that each work zone project should be completed within the predetermined time window. Eqs. (6.15) - (6.16) ensure that the maximum queue length per lane and average travel delay will not exceed the thresholds $Q_{\text{max}}$ and $T_a$, respectively.

### 6.4 Trial-and-error Solution Method

To solve the above total maintenance cost minimization model, the time window $[u_0, u_1]$ can be first discretized by using the $\Delta$-time interval, namely,

$$[u_0, u_1] = \left[ u_0, u_0 + \Delta t, \cdots, u_0 + \left( \frac{u_1 - u_0}{\Delta t} \right) \Delta t \right]$$  \hspace{1cm} (6.18)

where $\left( \frac{u_1 - u_0}{\Delta t} \right)$ is the maximum number of possible project start times in the time window $[u_0, u_1]$. The set of the possible project start times can be represented by

$$\Omega = \left\{ u_0 + j \times \Delta t \middle| j = 0, 1, 2, \cdots, \left( \frac{u_1 - u_0}{\Delta t} \right) \right\}$$ \hspace{1cm} (6.19)

According to the constraints shown by Eqs. (6.12)-(6.13), the number of subwork zones, $n$, should satisfy the following relation:

$$\left\lfloor \frac{L}{l_{\text{max}}} \right\rfloor \leq n \leq \left\lceil \frac{L}{l_{\text{min}}} \right\rceil$$ \hspace{1cm} (6.20)

Given any feasible $n$, all feasible project start times $y_1$ that satisfy the queue length and travel delay constraints can be found from the set of possible project start times $\Omega$. After enumerating all feasible $n$, all feasible subwork strategies can be obtained and the optimal subwork strategy can be determined by comparing the total maintenance costs caused by these feasible subwork zone strategies. Therefore, a trial-and-error solution method is designed as follows:
Step 1: (Initialization) Let the initial number of subwork zones \( n = \max \left\{ \left\lfloor \frac{L}{l_{\text{max}}} \right\rfloor, 1 \right\} \), the subwork zone length \( x_i = \frac{L}{n}, i = 1, 2, \ldots, n \), the initial project start time \( y_i = u_0, j = 0 \), the set of optimal solutions \( S^* = \emptyset \) and the initial minimum total maintenance cost \( C^* = +\infty \).

Step 2: (Calculation) Calculate average travel delay \( T_a(x, y) \), maximum queue length per lane \( Q_{\text{max}}(x, y) \) and the total maintenance cost \( C_{\text{MC}}(x, y) \) caused by the subwork zone operational strategy \((x, y)\).

Step 3: (Evaluation) If \( Q_{\text{max}}(x, y) > Q_{\text{max}} \) or \( T_a(x, y) > T_a \), go to Step 4; otherwise, perform the following manipulations:

Case 1: If \( C_{\text{MC}}(x, y) < C^* \), let \( C^* = C_{\text{MC}}(x, y) \), let the set of the optimal solutions \( S^* = \{(x, y)\} \);

Case 2: If \( C_{\text{MC}}(x, y) = C^* \), add the new solution \((x, y)\) into the set of optimal solutions, namely, \( S^* = S^* \cup \{(x, y)\} \);

Case 3: If \( C_{\text{MC}}(x, y) > C^* \), go to Step 4.

Step 4: (Project start time selection) If \( y_i \leq u_i - n \times \left( d_1 + d_2 \times \frac{L}{n} \right) \), let \( j = j + 1 \),

\[ y_i = u_0 + j \times \Delta t \]

and then go to Step 2; otherwise, let \( y_i = u_0, j = 0 \) and then go to Step 5.

Step 5: (stopping criterion) If \( n > \left\lfloor \frac{L}{l_{\text{min}}} \right\rfloor \), stop and output \( C^* \) and \( S^* \); otherwise, \( n = n + 1 \) and then go to Step 2.
6.5 Numerical Example

To assess the proposed model and the trial-and-error solution method, this chapter also employs the numerical example created from a short-term roadside tree-trimming project on four-lane two-way freeway in Singapore. The detailed description of this project has been presented in Section 5.6 and thus is not repeated. It is assumed that the total length of roadside trees to be trimmed in the project is 0.6 kilometers (e.g., \(L = 0.6 \text{ km}\)) and the time window is \([7:00, 18:00]\). Table 6.1 lists the input model parameters for the total maintenance cost minimization model.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Descriptions</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c_1)</td>
<td>Fixed cost of installing and removing devices for alerting/redirecting vehicles</td>
<td>S$45/h</td>
</tr>
<tr>
<td>(c_2)</td>
<td>Variable cost of installing and removing traffic control devices in the workspace</td>
<td>S$30/km·h</td>
</tr>
<tr>
<td>(c_3)</td>
<td>Fixed cost of traffic control devices for alerting/redirecting vehicles</td>
<td>S$50</td>
</tr>
<tr>
<td>(c_4)</td>
<td>Variable cost of traffic control devices in the workspace</td>
<td>S$70/km</td>
</tr>
<tr>
<td>(d_1)</td>
<td>Fixed setup time for a subwork zone</td>
<td>0.1h/zone</td>
</tr>
<tr>
<td>(d_2)</td>
<td>Variable work time per kilometer for a subwork zone</td>
<td>6h/zone</td>
</tr>
<tr>
<td>(l_{\text{min}})</td>
<td>The minimum length of a subwork zone restricted by traffic engineers</td>
<td>0.15km</td>
</tr>
<tr>
<td>(l_{\text{max}})</td>
<td>The maximum length of a subwork zone</td>
<td>0.60km</td>
</tr>
<tr>
<td>(C_{0+u+j})</td>
<td>Highway capacity during time interval ([u_0+j, u_0+j+1])</td>
<td>2550vph</td>
</tr>
<tr>
<td>(C_{w+u+j})</td>
<td>Work zone capacity during time interval ([u_0+j, u_0+j+1])</td>
<td>1383vph*</td>
</tr>
<tr>
<td>(V_{m+u+j})</td>
<td>Average speed of vehicles approaching work zone during time interval ([u_0+j, u_0+j+1])</td>
<td>40km/h</td>
</tr>
</tbody>
</table>

* work zone capacity estimated from the decision tree-based model in Chapter 4
Table 6.2 lists average hourly traffic (AHT) data and the corresponding average speed of vehicles approaching work zone, which is calculated using Eq. (5.49). The parameter $\Delta t$ is taken as 1 minute for discretizing the time window.

<table>
<thead>
<tr>
<th>Time interval $[u_0 + j, u_0 + j + 1)$</th>
<th>$f(u_0 + j)$</th>
<th>$V_{a,u_0+j}$ (km/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[7,8)</td>
<td>2361</td>
<td>85.4</td>
</tr>
<tr>
<td>[8,9)</td>
<td>1850</td>
<td>89.4</td>
</tr>
<tr>
<td>[9,10)</td>
<td>1390</td>
<td>89.8</td>
</tr>
<tr>
<td>[10,11)</td>
<td>1340</td>
<td>90.0</td>
</tr>
<tr>
<td>[11,12)</td>
<td>1735</td>
<td>89.8</td>
</tr>
<tr>
<td>[12,13)</td>
<td>1378</td>
<td>89.9</td>
</tr>
<tr>
<td>[13,14)</td>
<td>1596</td>
<td>89.8</td>
</tr>
<tr>
<td>[14,15)</td>
<td>1610</td>
<td>89.8</td>
</tr>
<tr>
<td>[15,16)</td>
<td>1725</td>
<td>89.8</td>
</tr>
<tr>
<td>[16,17)</td>
<td>1844</td>
<td>89.4</td>
</tr>
<tr>
<td>[17,18)</td>
<td>2120</td>
<td>87.6</td>
</tr>
</tbody>
</table>

6.5.1 Results and Discussions

According to Eq. (6.20), the entire work zone project can be divided into four subwork zones at most in this example. Figure 6.1 graphically depicts the relation between the total maintenance cost and the number of subwork zones. It shows that the total maintenance cost increases with the number of subwork zones. This is because the more the number of subwork zones, the more repeated setup costs would be caused. Note that the total maintenance cost is not affected by the project start time. If neither a travel delay constraint nor a queue length constraint were imposed by land...
transport authorities, the work zone contractor would prefer to close the entire lane for trimming roadside trees because the corresponding total maintenance cost is the least. This result is consistent with the findings in Chapter 5.

Figure 6.1 Total maintenance cost versus the number of subwork zones

Figure 6.2 Average travel delay versus project start time (y1)
Figure 6.2 displays the average travel delay with respect to different project start time choices. It can be observed that the minimum average travel delay varies with the subwork zone length, ranging from 2.6 min/veh to 3.2 min/veh. In addition to the subwork zone length, the project start time affects average travel delay as well. If the project is started earlier than 9:00 am, the corresponding average travel delay decreases with the project start time.

![Figure 6.2 Average travel delay with respect to project start time](image)

**Figure 6.2 Average travel delay with respect to project start time**

Figure 6.3 illustrates the effects of project start time on the maximum queue length. From the road user’s viewpoint, any time from 8:55 am to 9:00 am is the best start time of the project. This is because the maximum queue lengths associated with any project start time from 8:55 am to 9:00 am are the shortest, which are equal to 177 veh/lane. The influence of subwork zone length on the maximum queue length is

![Figure 6.3 Maximum queue length versus project start time](image)

**Figure 6.3 Maximum queue length versus project start time ($y_1$)**
shown in Figure 6.3 as well. If each subwork zone has a shorter length, the maximum queue length is larger when the roadside trimming project is started later than 9:20 am. This is because a short subwork zone length increases the number of repeated subwork zone setups. More repeated subwork zone setups result in a much later completion time so that more tree trimming work has to be implemented during the heavy traffic period (see in Table 6.3). Therefore, the maximum queue length is increased at the end time of the project.

<table>
<thead>
<tr>
<th>Time interval</th>
<th>$x_i$=0.60km (9:20-13:02)</th>
<th>$x_i$=0.30km (9:20-13:08)</th>
<th>$x_i$=0.20km (9:20-13:14)</th>
<th>$x_i$=0.15km (9:20-13:20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queue length</td>
<td>Queue length</td>
<td>Queue length</td>
<td>Queue length</td>
<td>Queue length</td>
</tr>
<tr>
<td>(veh/lane)</td>
<td>(veh/lane)</td>
<td>(veh/lane)</td>
<td>(veh/lane)</td>
<td>(veh/lane)</td>
</tr>
<tr>
<td>9:00-10:00</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10:00-11:00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11:00-12:00</td>
<td>178</td>
<td>178</td>
<td>178</td>
<td>178</td>
</tr>
<tr>
<td>12:00-13:00</td>
<td>177</td>
<td>177</td>
<td>177</td>
<td>177</td>
</tr>
<tr>
<td>13:00-14:00</td>
<td>178</td>
<td>181</td>
<td>202</td>
<td>213</td>
</tr>
</tbody>
</table>

$Q_{\text{max}}(x,y)$ (veh/lane) 178 181 202 213

This chapter considers the following three scenarios with different thresholds $T_a$ and $Q_{\text{max}}$:

Scenario 1: $T_a^{(1)} = 3.5$ min/veh and $Q_{\text{max}}^{(1)} = 200$ veh/lane;

Scenario 2: $T_a^{(2)} = 3.0$ min/veh and $Q_{\text{max}}^{(2)} = 180$ veh/lane;

Scenario 3: $T_a^{(3)} = 2.5$ min/veh and $Q_{\text{max}}^{(3)} = 160$ veh/lane.
The two thresholds of these three scenarios are respectively shown in Figure 6.2 and 6.3. In Scenario 1, it can be seen that the average travel delay and maximum queue length caused by the single work zone strategy \( x_i = 0.60 \text{ km} \) are respectively less than \( T_a^{(1)} \) and \( Q_{\text{max}}^{(1)} \) when the project starts at any time between 8:59 am and 9:31 am. As aforementioned earlier, a single work zone strategy that implies the longest subwork zone length is preferred by the roadside tree-trimming contractor if a feasible project start time can be found to meet the queue length and travel delay thresholds. Therefore, the contractors will close the entire lane for trimming roadside trees and starts the project at any time between 8:59 am and 9:31 am in this scenario. The resulting total maintenance cost is \$325.1. Figure 6.2 and 6.3 also indicate that dividing the entire work zone into two subwork zones with the uniform length of 0.30 km is the optimal work strategy in the second scenario. The optimum project start time is found to be any time between 8:59 am and 9:13 am. Since the average travel delay and maximum queue length with respect to any work strategy cannot meet the travel delay and maximum queue length requirements, there is no feasible subwork strategy for the contractor in the third scenario. Therefore, it can be concluded that the average travel delay and queue length thresholds have significant impacts on the optimal subwork zone length and project start time. It should be pointed out that the above analysis also illustrates the process of the trial-and-error method to determine the optimal solutions.

### 6.5.2 Model Comparison

This section will compare the optimal solutions from the proposed total maintenance cost model and the total work zone cost minimization model (5.25)
-(5.31) in Scenario 2. The trial-and-error method proposed in this chapter is also capable of solving the total work zone cost minimization model (5.25)-(5.31).

| Table 6.4 Comparisons of results from the total work zone cost model and the total maintenance cost model |
|-------------------------------------------------|-------------------------------------------------|
| Model (6.11)-(6.17) | Model (5.25)-(5.31)* |
| Queue length threshold | $Q_{max}^2 = 180$ veh/lane | No |
| Travel delay threshold | $T_d^2 = 3.0$ min/veh | No |
| Maximum queue length | 177 veh/lane | 184 veh/lane |
| Average travel delay | 2.9 min/veh | 2.7 min/veh |
| Total travel delay | 265.6~268.4 vehicle-hours | 259.6 vehicle-hours |
| Total maintenance cost | S$347.2 | S$390.9 |
| Total work zone cost | S$3535.1~3568.2 | S$3506.2 |
| Optimum number of subwork zones ($n$) | 2 | 3 |
| Optimal subwork zone length ($x_i$) | 0.30 km | 0.20 km |
| Best project start time ($y_1$) | Any time between 8:59 am and 9:13 am | 9:08 am |

$c_{vh}$ is assumed as 12 S$/vehicle-hour for the total work zone cost model(5.25)-(5.31).

Table 6.4 gives two different optimal solutions, which are generated by the presented total maintenance cost minimization model and the total work zone cost minimization model, respectively. The optimal solution from the proposed total maintenance cost model is to adopt the uniform subwork zone length of 0.30 km and start the project at any time between 8:59 am and 9:13 am. The optimal solution yielded by the total work zone cost model (5.25)-(5.31) is to divide the entire work zone into three subwork zones with the uniform length of 0.20 km at the project start time of 9:10 am, which results in the minimum total work zone cost of S$3506.2. This
total work zone cost is 0.8~1.7% lower than the corresponding value from the proposed model. Moreover, the total maintenance cost of S$390.9 is about 12.6% higher, as compared with corresponding value from the total maintenance cost minimization model.

However, the optimal solution of total work zone cost minimization model cannot meet the queue length threshold because the resulted maximum queue length is 184 veh/lane, larger than the queue length threshold $Q_{\text{max}}^{(2)} (=180$ veh/lane). In this situation, the contractor may require to pay additional lane rental charges or penalties imposed by land transport authorities for not complying with the queue length and travel delay constraints written in the contract.

### 6.6 Summary

This chapter has presented a new perspective to determine the optimal subwork zone operational strategy by minimizing the total maintenance cost subject to the queue length and travel delay constraints imposed by the land transport authorities. An efficient trial-and-error method has been proposed to find all the optimal solutions for the proposed model.

The numerical results have shown the availability and merits of the proposed model and solution method. They also demonstrated that the optimal subwork zone operational strategy from the contractor’s perspective depends on the average travel delay and queue length constraints. Using the proposed model, the contractor will be able quickly to determine the best operational strategy that minimizes the total maintenance cost as well as complies with the queue length and travel delay constraints.
CHAPTER 7 WORK ZONE REAR-END CRASH RISK ASSESSMENT

7.1 Introduction

The presence of a work zone increases traffic conflicts, which could increase the rear-end crash potential. Previous studies reported that the crash rate in a work zone is higher than that in non-work zone (Wang et al., 1996; Roupail et al., 1988; Khattak et al., 2002). Further, Pigman and Agent (1990) found that about 80% of all work zone crashes occur in the work zone activity area according to an analysis of crashes in Kentucky. Based on 1484 crashes occurred in Virginia between 1996 and 1999, Garber and Zhao (2002) found that the activity area is the predominant location for work zone crashes regardless of road type and that rear-end crashes are the predominant type of crash. Although an optimal subwork zone operational strategy from the total maintenance cost minimization model could reduce work zone maintenance cost, it may lead to high rear-end crash due to the queues. Therefore, concerns should be also addressed on the analysis of rear-end crash risk at work zones under a given subwork zone operational strategy. Hereafter, the rear-end crash risk is referred to be the probability that a vehicle is involved in a rear-end crash when it travels across the entire work zone activity area.

A number of studies have been conducted on the rear-end crash analysis. In these studies, historical accident data are used to identify the causal factors and analyze the injury severity using statistical analysis techniques. In general, this is an
efficient method of analyzing crash risk. However, it may sometimes result in inaccurate or biased results if the historical data have poor quality and reliability when the traffic police wrongly record accidents (Kamalasudhan et al., 2002). There is also a possibility that no historical accident data will be available, especially for a newly-built road or a newly-proposed work zone. Fortunately, the advanced sensor and video technology allows us to estimate rear-end crash risk by using the available traffic data from the field. Recent studies have applied surrogate safety measures (SSM) to measure safety based on the vehicle trajectory data, which provide adequate evidence that the models based on traffic data can yield reliable and accurate risk estimates.

The purpose of this chapter is to assess the rear-end crash risk in the work zone activity area based on the traffic data collected from the field. In addition, this chapter also investigates the effects of contributing factors (e.g., lane position and heavy vehicle percentage) that have not been fully examined in previous studies. According to Cunto and Saccomanno (2008), the deceleration rate to avoid the crash (DRAC) can be regarded as an effective index to assess the vehicle rear-end crash risk. Hence, DRAC is adopted in measuring the probability of a vehicle involving in a rear-end crash when traveling through the work zone activity area.

7.2 Data

7.2.1 Work Zone Site Description

To evaluate the rear-end crash risk in the work zone activity area, a field
survey on an arterial work zone site and an expressway work zone site in Singapore is conducted. Figure 7.1 shows the layouts of the above two work zone sites.

![Figure 7.1 Work zone sites for field data collection](image)

(a). Expressway work zone

(b). Arterial work zone

**Figure 7.1 Work zone sites for field data collection**

The arterial work zone site is located in the arterial road of Ang MoKio Avenue 3. In this site, two lanes are opened while the fast lane is closed for maintenance activities. The length of this arterial work zone site is about 300 m. The expressway work zone site, located in the four-lane one-way Central Expressway
(CTE), causes the fast lane reduction for construction activities, as shown in Figure 7.1. The length of the expressway work zone site is about 400 m. At both work zone sites, the traffic data during the peak and non-peak periods are recorded.

### 7.2.2 Data Collection

A video camera can provide a continuous monitoring of traffic flows at work zones. Compared with the other data collection methods such as the application of manual equipment, this method can reduce stochastic errors because all vehicle trajectory data can be obtained by replaying the collected videotape. Hence, a video camera is used to record work zone traffic.

For the expressway and arterial work zones, traffic data are extracted from the collected videotapes on a lane-by-lane basis. Each lane traffic data is further split into 15-minute time intervals. For the sake of presentation, one 15-minute time interval of lane traffic is defined as one work zone scenario.

**Traffic data**

For each vehicle, its type is first determined, denoted by $V_{type}$. In this study, vehicles are grouped into two categories: i) car and ii) truck. Hereafter, a car is referred to as a vehicle having less than two axles or having less than four wheels (e.g., private car, cabs and minivan). A truck is defined as a vehicle having not less than two axles or not less than four wheels (e.g., truck, bus). Note that motorcycles are not considered in this study.

After identifying the type of a vehicle, its speed is then measured using the
Premier Pro CS3 software. This software can display 30 frames per second and the error is 0.03 s. Due to the analyst’s visual judgment error for vehicle positions, the total possible error could be up to 0.1 s (Bonneson and FITTS, 1995). The speed of a particular vehicle \( n \), denoted by \( v_n \), is determined by calculating the time taken by this vehicle to cover a two lane-markers’ distance in the video. As pointed out by Strong et al (2003), this measurement method could yield data of comparable quality to the radar-speed measurement method, which produces the potential error of ±0.8 km/h.

In addition to the Premier Pro CS3 software, two traffic data collectors (TDC) are used to assist in recording the time-headway \( h \) of two vehicles \( n \) and \( n-1 \). According to the study of Vogel (2003), the gap size \( d_{n-1}^{n} \) between the follower vehicle \( n \) and the lead vehicle \( n-1 \) can be calculated as \( d_{n-1}^{n} = v_n h - l_{n-1} \). Here, \( l_{n-1} \) is the length of the lead vehicle \( n-1 \).

Finally, the observed vehicle trajectory data are used to determine the lane traffic volume, lane traffic speed and heavy vehicle percentage by means of the conventional method at each work zone scenario. The equivalent hourly lane traffic flow rate is calculated as four times of the 15-minute volume.

**Vehicle rear-end crash risk**

A number of SSM including the time to collision (TTC), the DRAC and the post encroachment time (PET) have been applied to measure road safety and examine the relevant contributing factors. Minderhoud and Bovy (2001) used the time integrated time-to-collision (TIT) and the time exposed time-to-collision (TET) based on the TTC notion to analyze road traffic safety. Alexander et al. (2002) applied the
gap size to measure the safety of unsignalized junction and examined the relationship between gap size and contributing factors. Archer (2005) has explicitly recognized the superiority of $DRAC$ as a safety measure indicator. This indicator reflects the follower vehicle deceleration required to come to a timely stop or attain the corresponding leader vehicle speed in order to avoid a rear-end crash.

According to the $DRAC$ definitions in previous studies, the deceleration rate of the vehicle $n$ to avoid the crash with the vehicle $n-1$, denoted by $DRAC_{n-1}^{n}$, can be calculated as

$$DRAC_{n-1}^{n} = \begin{cases} \frac{(v_n - v_{n-1})^2}{d_{n-1}^n}, & \forall v_n > v_{n-1} \\ 0, & \forall v_n \leq v_{n-1} \end{cases}$$  \quad (7.1)

where $v_{n-1}$ is the speed of the lead vehicle $n-1$, $d_{n-1}^n$ is the gap size between the follower vehicle $n$ and the lead vehicle $n-1$. The deceleration rate of the vehicle $n+1$ to avoid the crash with the vehicle $n$, denoted by $DRAC_{n+1}^n$, is also estimated using Eq. (7.1).

However, researchers argued that conventional $DRAC$ fails to accurately identify the potential crash situations because it does not consider the vehicle braking capability for prevailing road conditions. For instance, it is expected that a small value of $DRAC$ under the wet pavement conditions is more dangerous than the same value under the dry pavement conditions because the vehicle traveling in the wet pavement conditions has a lower braking capability due to the reduced road friction coefficient.

Therefore, the $DRAC$ is inadequate to describe road safety.

Taking into account the vehicle braking capability, Cunto and Saccomanno (2007) proposed a method to determine the crash risk based on the calculated $DRAC$. Similar to their study, this chapter also calculates the rear-end crash risk between two
vehicles as the probability that a given DRAC exceeds its maximum available deceleration rate (MADR). Since MADR depends on the factors such as pavement conditions (dry/wet/snow), vehicle weight, tire and braking system, it should be considered to be a random variable. In this study, the MADR is assumed to follow a truncated normal distribution and the parameters for MADR are listed in Table 7.1.

| Table 7.1 Truncated normal distribution parameters of MADR at dry pavement conditions |
|-----------------------------------------------|-------------------|
| MADR distribution parameters             | Car (m/s²)   | Truck (m/s²) |
| Mean (m/s²)                             | 8.45<sup>a</sup> | 6.82<sup>b</sup> |
| Standard deviation (m/s²)               | 1.40          | 1.40          |
| Upper limit (m/s²)                      | 12.68         | 10.05         |
| Lower limit (m/s²)                      | 1.23          | 0.60          |

<sup>a</sup> source from Cunto and Saccomanno (2008)
<sup>b</sup> source from AASHTO (2004)

In reality, the vehicle <i>n</i> has the probability of crash with its lead vehicle <i>n-1</i> as well as the probability of being collided by its lag vehicle <i>n+1</i>. Therefore, the probability of rear-end crash for the vehicle <i>n</i> at time <i>t</i>, denoted by <i>p_{n,t,n}</i>, is calculated as the sum of the probabilities of colliding with the vehicles <i>n-1</i> and <i>n+1</i>, shown as follows:

\[
p_{n,t} = p\left( DRAC^{n-1}_{n,j} > MADR_{n,t} \right) + p\left( DRAC^{n}_{n+1,j} > MADR_{n+1,t} \right)
\]

(7.2)

where \( p\left( DRAC^{n-1}_{n,j} > MADR_{n,t} \right) \) is the probability of a rear-end crash between vehicles <i>n</i> and <i>n-1</i> at time <i>t</i>, and \( p\left( DRAC^{n}_{n+1,j} > MADR_{n+1,t} \right) \) is the probability of a rear-end crash between vehicles <i>n</i> and <i>n+1</i> at time <i>t</i>.

The average risk of a vehicle of type <i>i</i> involving in a rear-end crash at each work zone scenario can thus be calculated by: 
where \( T_n \) is the time taken by the vehicle \( n \) traveling through the entire work zone activity area, \( \Delta T \) is the observation time interval, \( N \) is the number of vehicles passing through the work zone, \( I(V_{type,n} = i) \) is a binary variable. \( I(V_{type,n} = i) \) is equal to 1 if the type of vehicle \( n \) belongs to the \( i \)th type; otherwise it is equal to zero.

### 7.3 Modeling Rear-end Crash Risk at Arterial & Expressway Work Zones

This chapter takes into account the following candidate variables, which may affect the vehicle’s rear-end crash risk at work zones:

1. The lane traffic flow rate, \( f(vphl) \).
2. The lane traffic speed, \( V \) (km/h).
3. The heavy vehicle percentage, \( hv \).
4. The lane position, \( L_{pos} \). In this chapter, the value of \( L_{pos} \) is determined depending on the proximity of the lane to work zone. For instance, \( L_{pos} \) for the lane which is closest to the work zone is set to 1.
5. An indicator variable for vehicle types, \( V_{type} \). \( V_{type} = 1 \) for a car and \( V_{type} = 2 \) for a truck.
6. An indicator variable for work zone types, \( W_{type} \). \( W_{type} = 1 \) for an expressway work zone and \( W_{type} = 2 \) for an arterial work zone.

To discern the meaningful relationships between rear-end crash risk and its
contributing factors, two rear-end crash risk models are first formulated based on the data collected from the expressway and arterial work zones, respectively. These two models do not intend to examine the effect of work zone type on the rear-end crash risk. Hence, the first five candidate variables mentioned above are used to determine forms of the two models. Based on the combination of data from both work zone types, the third rear-end crash risk model is then developed to provide the overall information on factors affecting rear-end crash risk. All the six candidate variables are thus taken into account in the third model.

In general, each rear-end crash risk model can be expressed by a generalized linear form:

$$\ln(R) = a_0 + \sum_{i=1}^{N_a} (a_i \times x_i) + \epsilon$$

(7.4)

where $R$ is the rear-end crash risk, $a_0, a_i, i=1,\ldots,N_a$ are the coefficients to be estimated, $\epsilon$ is the normally distributed random error, and $x_i, i=1,2,\ldots,N_a$ are the explanatory variables associated with the above candidate variables, $N_a$ is the number of explanatory variables adopted by the model.

It is assumed that each of the above candidate variables has three possible relations with respect to its relevant contributing factor: identical, logarithm and reciprocal. For example, the possible form of the lane position $L_{pos}$ may be $L_{pos}$, $\ln(L_{pos})$ or $1/L_{pos}$. Therefore, the maximum number of explanatory variables is $3 \times 5 = 15$ for the first two models and $3 \times 6 = 18$ for the third model.

The stepwise regression analysis method is adopted to determine the model form shown in Eq.(7.4). In the stepwise regression analysis, the stepwise selection method is adopted to select explanatory variables and determine the coefficients associated with the selected explanatory variables for the crash risk model. This
method is a modification of the forward selection technique and differs in that variables already in the model do not necessarily stay there. As in the forward selection method, variables are added one by one to the model, and the $F$ statistic for a variable to be added must be significant at the level of the preset entry value, $\alpha_{\text{entry}}$. After a variable is added, however, the stepwise method looks at all the variables already included in the model and deletes any variable with the level of significance greater than the preset remove value, $\alpha_{\text{remove}}$. Only after this check is made and the necessary deletions accomplished can another variable be added to the model. The stepwise process ends when none of the variables outside the model is significant at the level of $\alpha_{\text{entry}}$ and every variable included in the model is significant at the level of $\alpha_{\text{remove}}$.

7.4 Analysis of Results

7.4.1 ANOVA Analysis of Rear-end Crash Risk

A total of 9,350 sets of vehicle trajectory data are collected, including 6,548 sets from the expressway work zone and 2,802 sets from the arterial work zone. Figure 7.2 illustrates the relationship between the estimated $DRAC$ and its contributing factors. According to the figure, it can be seen that the $DRAC$ most likely to increase with the lane traffic flow and heavy vehicle percentage while decrease with the lane position and work zone type.
Based on the collected data, a one-way ANOVA analysis is carried out to test whether the mean of rear-end crash risk is statistically different from the vehicle types and lane positions at work zones. Table 7.2 gives the ANOVA test results for the expressway work zone. According to the table, it can be found that the mean rear-end crash risk of a car is less than that of a truck at the expressway work zone. However, it does not show that the means of rear-end crash risk at different vehicle types are statistically different at the significance level of 0.05. The means of rear-end crash risk at the expressway work zone are significantly different at different lane positions.
Table 7.2 Statistical comparisons of rear-end crash risk for different vehicle types and lane positions at the expressway work zone (sample size=6548)

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>Stdev.</th>
<th>$F$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vehicle type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>1.59E-05</td>
<td>1.31E-04</td>
<td>0.54</td>
<td>0.464</td>
</tr>
<tr>
<td>Truck</td>
<td>1.96E-05</td>
<td>1.41E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lane position</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane 1</td>
<td>2.87E-05</td>
<td>1.99E-04</td>
<td>7.00</td>
<td>0.001</td>
</tr>
<tr>
<td>Lane 2</td>
<td>1.01E-05</td>
<td>7.10E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane 3</td>
<td>1.09E-05</td>
<td>8.14E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All vehicles</td>
<td>1.71E-05</td>
<td>1.34E-04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The rear-end crash risk of different vehicle types at the arterial expressway work zone are found to be statistically significant at the level of 0.061, as shown in Table 7.3. The mean of rear-end crash risk in Lane 1 ($L_{pos}=1$) is also found to be statistically larger than that in Lane 2 ($L_{pos}=2$). Table 7.4 shows the summary of the one-way ANOVA for the combination of data from the expressway and arterial work zones. From the table, it can be clearly seen that the means of rear-end crash risk for the expressway and arterial work zones are statistically different for different vehicle types and different lane positions at the significance level of 0.05. Comparing with Table 7.2-7.3, it can be observed that the mean of rear-end crash risk at the expressway work zone is much larger than that at the arterial work zone. In addition, according to Table 7.2-7.4, it can be seen that the standard deviation of the rear-end crash risk is substantially higher than the mean, which confirms the appropriateness of using the logarithmic transformation of the target variable $R$ for building the rear-end
crash risk models.

Table 7.3 Statistical results of rear-end crash risk for different vehicle types and lane positions at the arterial work zone (sample size=2802)

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>Stdev</th>
<th>$F$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>2.14E-06</td>
<td>4.21E-05</td>
<td>3.50</td>
<td>0.061</td>
</tr>
<tr>
<td>Truck</td>
<td>5.69E-06</td>
<td>5.25E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane position</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane 1</td>
<td>4.50E-06</td>
<td>6.15E-05</td>
<td>4.85</td>
<td>0.028</td>
</tr>
<tr>
<td>Lane 2</td>
<td>8.50E-07</td>
<td>7.47E-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All vehicles</td>
<td>2.67E-06</td>
<td>4.38E-05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.4 Statistical results of rear-end crash risk ($R$) for different vehicle types and lane positions at the expressway and arterial work zones (sample size=9350)

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>Stdev</th>
<th>$F$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>8.80E-06</td>
<td>9.59E-05</td>
<td>5.00</td>
<td>0.025</td>
</tr>
<tr>
<td>Truck</td>
<td>1.56E-05</td>
<td>1.23E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane position</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane 1</td>
<td>1.56E-05</td>
<td>1.43E-04</td>
<td>6.68</td>
<td>0.001</td>
</tr>
<tr>
<td>Lane 2</td>
<td>5.20E-06</td>
<td>4.89E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane 3</td>
<td>1.09E-05</td>
<td>8.14E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All vehicles</td>
<td>1.28E-05</td>
<td>1.15E-04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7.4.2 Analysis of Rear-end Crash Risk Model Results

According to the lane position, vehicle type and time interval length, the collected 9,350 sets of data are further categorized into 68 work zone scenarios, including 36 work zone scenarios for the expressway work zone and 32 work zone scenarios for the arterial work zone. Each work zone scenario consists of the information regarding the hourly lane traffic flow rate, the traffic speed, the lane position, the vehicle type and the average rear-end crash risk. The regression procedure (Proc REG) provided by the SAS (2008) is implemented to perform the stepwise regression analysis to select explanatory variables and determine the coefficients associated with the selected explanatory variables for the rear-end crash risk models. The parameters of $\alpha_{\text{entry}}$ and $\alpha_{\text{remove}}$ are both taken as 0.10 in this study.

Model 1: rear-end crash risk model for the expressway work zone

Table 7.5 shows the explanatory variable selection results by using the stepwise selection method. It can be seen that the explanatory variable $\ln(L_{\text{pos}})$ is always selected during the stepwise selection process. The regression procedure outputs the highest $R^2$ of 84.7% when four explanatory variables, including $\ln(L_{\text{pos}})$, $V_{\text{type}}$, $1/hv$ and $1/f$, are selected for building the rear-end crash risk model at expressway work zones (Model 1). Therefore, $N_a$ in Eq. (7.4) is equal to four for Model 1.
Table 7.5 Stepwise selection results of the explanatory variables for Model 1

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td></td>
<td></td>
<td></td>
<td>$\sqrt{\quad}$</td>
</tr>
<tr>
<td>$S$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$hv$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{type}$</td>
<td>$\sqrt{\quad}$</td>
<td>$\sqrt{\quad}$</td>
<td>$\sqrt{\quad}$</td>
<td>$\sqrt{\quad}$</td>
</tr>
<tr>
<td>$L_{pos}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(f)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(S)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(hv)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(V_{type})$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(L_{pos})$</td>
<td>$\sqrt{\quad}$</td>
<td>$\sqrt{\quad}$</td>
<td>$\sqrt{\quad}$</td>
<td>$\sqrt{\quad}$</td>
</tr>
<tr>
<td>$1/f$</td>
<td></td>
<td></td>
<td></td>
<td>$\sqrt{\quad}$</td>
</tr>
<tr>
<td>$1/S$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1/hv$</td>
<td></td>
<td></td>
<td>$\sqrt{\quad}$</td>
<td>$\sqrt{\quad}$</td>
</tr>
<tr>
<td>$1/V_{type}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1/L_{pos}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Model $R^2$       | 69.4% | 79.9% | 83.5% | **84.7%** |

Note: the symbol “$\sqrt{\quad}$” represents that the variable is adopted by the model.

Table 7.6 gives the coefficients of the selected explanatory variables and the corresponding statistical test results for Model 1. According to the correlation analysis results, it can be seen that only the lane position has a little high correlation with the lane traffic flow (0.353) and heavy vehicle percentage (0.452). According to Table 7.6, it can be also seen that Kolmogorov-Smirnov and Anderson-Darling tests’ $p$-values on the error term are all larger than 0.15, suggesting that there is no evidence that the data do not follow a normal distribution. This important finding confirms the
feasibility and reasonability of generalized linear model form shown by Eq. (7.4).

Table 7.6 Statistical results of the rear-end crash risk model for the expressway work zone (sample size=36)

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>$F$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-8.884</td>
<td>0.638</td>
<td>193.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$V_{type}$</td>
<td>0.957</td>
<td>0.176</td>
<td>29.59</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$1/hv$</td>
<td>-0.383</td>
<td>0.314</td>
<td>11.98</td>
<td>0.001</td>
</tr>
<tr>
<td>$1/f$</td>
<td>-1021</td>
<td>562.7</td>
<td>3.29</td>
<td>0.077</td>
</tr>
<tr>
<td>$\ln(L_{pos})$</td>
<td>-1.714</td>
<td>0.363</td>
<td>22.32</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Goodness-of-Fit Tests for Normal Distribution*

<table>
<thead>
<tr>
<th></th>
<th>$p$-value</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>&gt;0.150</td>
<td></td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>0.205</td>
<td></td>
</tr>
</tbody>
</table>

Correlation analysis

<table>
<thead>
<tr>
<th></th>
<th>$V_{type}$</th>
<th>$1/hv$</th>
<th>$1/f$</th>
<th>$\ln(L_{pos})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{type}$</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$1/hv$</td>
<td>1.000</td>
<td>0.025</td>
<td>0.452</td>
<td></td>
</tr>
<tr>
<td>$1/f$</td>
<td>1.000</td>
<td>0.353</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(L_{pos})$</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Tests on the random error term $\varepsilon$.

Therefore, Model 1 can be written as follows

$$R = (L_{pos})^{-1.714} \exp(0.957V_{type})\exp(-0.383/hv)\exp(-1021/f)\exp(-8.884) \quad (7.5)$$

The relative effects of variables on the rear-end crash risk at the expressway work zone can be calculated from the model shown by Eq.(7.5). The estimated log-transformed coefficient of $L_{pos}$ is negative and statistically significant at the 0.05 level, indicating that the lane closer to the work zone is strongly associated with a
higher rear-end crash risk. More specifically, the rear-end crash risk in Lane 1 ($L_{pos}=1$) is about three ($\approx 1/2^{1.714}$) times of that in Lane 2 ($L_{pos}=2$). The estimated coefficient associated with the indicator variable of vehicle types ($V_{type}=1$ for a car and $V_{type}=2$ for a truck) is positive. It simply implies that the rear-end crash risk of a car traveling at the expressway work zone is lower than that of a truck. This finding is not surprising because the truck has a less braking capability.

According to the model results, the reciprocal-transformed coefficients of the heavy vehicle percentage and the lane traffic flow rate are negative. This suggests that the rear-end crash risk increases with the lane traffic flow rate and the heavy vehicle percentage. However, the model (Table 7.6) also shows that the exposure term of variable $1/f$ is not statistically significant at the level of 0.10. One possible reason for this is that the exposure effects are already captured individually by the lane position and heavy vehicle percentage.

Model 2: rear-end crash risk model for the arterial work zone

Table 7.7 shows the factors that influence the rear-end crash risk at the arterial work zone and the corresponding statistical test results. It can be found that only the lane traffic flow has a little high correlation with the lane position (0.555). The normality test results in Table 7.7 confirm that the random error term $\epsilon$ follows a normal distribution. The rear-end crash risk model at the arterial work zone (Model 2) can be written as:

$$R = (L_{pos})^{-1.231} \exp(0.661V_{type}) \exp(2.246h) \exp(-347.9/f) \exp(-14.93) \quad (7.6)$$
Table 7.7 Statistical results of the rear-end crash risk model for the arterial work zone (sample size=32)

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>$F$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-14.93</td>
<td>0.403</td>
<td>1371</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$V_{type}$</td>
<td>0.661</td>
<td>0.214</td>
<td>9.55</td>
<td>0.004</td>
</tr>
<tr>
<td>$h_v$</td>
<td>2.246</td>
<td>0.849</td>
<td>6.99</td>
<td>0.012</td>
</tr>
<tr>
<td>$1/f$</td>
<td>-347.9</td>
<td>114.6</td>
<td>9.21</td>
<td>0.004</td>
</tr>
<tr>
<td>$\ln(L_{pos})$</td>
<td>-1.231</td>
<td>0.380</td>
<td>10.48</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Goodness-of-Fit Tests for Normal Distribution

Kolmogorov-Smirnov ($p$-value) $>0.150$

Anderson-Darling ($p$-value) 0.345

Correlation analysis

<table>
<thead>
<tr>
<th>$V_{type}$</th>
<th>$h_v$</th>
<th>$1/f$</th>
<th>$\ln(L_{pos})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{type}$</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$h_v$</td>
<td>1.000</td>
<td>0.344</td>
<td>0.121</td>
</tr>
<tr>
<td>$1/f$</td>
<td>1.000</td>
<td>0.555</td>
<td></td>
</tr>
<tr>
<td>$\ln(L_{pos})$</td>
<td></td>
<td></td>
<td>1.000</td>
</tr>
</tbody>
</table>

The $R^2$ of Model 2 shown by Eq. (7.6) is 70.2%. Similar to Model 1, the sign of the estimated log-transformed coefficient of $L_{pos}$ is also negative in Model 2. However, in comparison with the expressway work zone, the arterial work zone has the smaller effects of lane positions on the rear-end crash risk. For example, the rear-end crash risk in Lane 1 ($L_{pos}$=1) is about twice ($\approx 1/2^{-1.231}$) of that in Lane 2 ($L_{pos}$=2) for the arterial work zone, less than three times for the expressway work zones. Similarly, the relative effect of vehicle types on the rear-end crash risk is $\exp(0.661)=1.94$ for the arterial work zone, less than 2.60 ($\exp(0.957)$) for the
expressway work zone. It shows that the truck has a marginal stronger effect on the increase of rear-end crash risk at the expressway work zone compared with the increase of rear-end crash risk at the arterial work zone.

As expected, the heavy vehicle percentage shows a positive coefficient to the rear-end crash risk. In general, 0.01 unit increase of heavy vehicle percentage will result in a 2.3% \( (e^{0.01} - 1) \) increase of rear-end crash risk. According to Table 7.7 and Eq. (7.6), it can be found that rear-end crash risk increases with the lane traffic flow rate. This result is not surprising because the average gap between two vehicles becomes shorter as the number of vehicles per lane increases, which eventually increases the likelihood of crashes. This finding is consistent with the previous studies, which focused on evaluating work zone crash rates (Khattak et al., 2002).

Model 3: combined rear-end crash risk model for the expressway and arterial work zones

Table 7.8 presents the results of the combined rear-end crash risk model (Model 3) estimated for both work zone types. The combined model has a relatively higher model performance \( (R^2=88.2\%) \), compared with Model 1 and Model 2. In Model 3, five explanatory variables are selected and the corresponding coefficients are shown in Table 7.8. According to the table, it can be clearly seen that the coefficients associated with the selected explanatory variables are all statistically significant at the level of 0.05. Hence, the combined rear-end crash risk model (Model 3) can be written as:
\[ R = \exp(-0.761L_{\text{pos}}) \exp(0.861V_{\text{type}}) \exp(2.919hv) \exp(-331/f) \times \exp(-3.380W_{\text{type}}) \exp(-7.848) \]  

(7.7)

Table 7.8 Summary of the combined rear-end crash risk model for the expressway and arterial work zones (sample size=68)

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>( F )</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-7.848</td>
<td>0.919</td>
<td>72.83</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>( V_{\text{type}} )</td>
<td>0.816</td>
<td>0.202</td>
<td>16.29</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>( hv )</td>
<td>2.919</td>
<td>0.673</td>
<td>18.83</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>( 1/f )</td>
<td>-331.0</td>
<td>138.0</td>
<td>5.75</td>
<td>0.021</td>
</tr>
<tr>
<td>( L_{\text{pos}} )</td>
<td>-0.761</td>
<td>0.203</td>
<td>14.10</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>( W_{\text{type}} )</td>
<td>-3.380</td>
<td>0.363</td>
<td>86.68</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Goodness-of-Fit Tests for Normal Distribution

Kolmogorov-Smirnov \((p\)-value\) 0.207

Anderson-Darling \((p\)-value\) 0.242

Correlation analysis

<table>
<thead>
<tr>
<th></th>
<th>( V_{\text{type}} )</th>
<th>( hv )</th>
<th>( 1/f )</th>
<th>( L_{\text{pos}} )</th>
<th>( W_{\text{type}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{\text{type}} )</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( hv )</td>
<td></td>
<td>1.000</td>
<td>-0.077</td>
<td>-0.298</td>
<td>-0.311</td>
</tr>
<tr>
<td>( 1/f )</td>
<td></td>
<td></td>
<td>1.000</td>
<td>0.376</td>
<td>0.447</td>
</tr>
<tr>
<td>( L_{\text{pos}} )</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>-0.056</td>
</tr>
<tr>
<td>( W_{\text{type}} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
</tr>
</tbody>
</table>

Similar to Model 1 and Model 2, the rear-end crash risk in Model 3 also decreases with the lane position and increases with the vehicle type, heavy vehicle percentage and lane traffic flow rate.

In contrast to Model 1 and 2, the combined rear-end crash risk model can
show the effect of work zone type on the rear-end crash risk. In Table 7.8, the coefficient of work zone type indicator ($W_{\text{type}}$ =1 for an expressway work zone and $W_{\text{type}}$ =2 for an arterial work zone) is negative and statistically significant at the level of 0.05. It implies that the rear-end crash risk at the expressway work zone is significantly larger than that at the arterial work zone, which is consistent with the ANOVA test result in this study and the findings from Garber and Zhao (2002). One reason for the higher rear-end crash risk at the expressway work zone is that the lane traffic flow rate and vehicle’s traveling speed at the expressway work zone are larger than at the arterial work zone. In general, the larger traveling speed will increase the likelihood of a rear-end crash. Another reason may be that the expressway work zone for construction activities is actually a long-term work zone. As deemed by many researchers (Rouphail et al., 1988; Ullman and Krammes 1990), there will be considerable increases of crash potential in the long-term expressway work zones.

7.5 Summary

To address concerns on the rear-end crash risk at work zones, this chapter estimated the work zone rear-end crash risk and examined the effects of contributing factors based on the available work zone traffic data. The DRAC was used in measuring the risk of a vehicle involving in a rear-end crash when traveling through the work zone activity area. Three rear-end crash risk models were developed in this study. In the first two models, the rear-end crash risk was respectively formulated as a
function of the contributing factors, including: i) lane traffic flow rate; ii) lane traffic speed; iii) heavy vehicle percentage, iv) lane position and v) vehicle type. The third model was the combination of the first two models by adding a new contribution factor—work zone type. A stepwise regression method was adopted to select explanatory variables and determine the coefficients associated the selected variables for the models. The traffic data collected from the expressway and arterial work zones in Singapore were finally utilized to calibrate the three rear-end crash risk models.

The one-way ANOVA test results showed that the rear-end crash risk is statistically different from the lane positions at the expressway and arterial work zones. The rear-end crash risk at the arterial work zone is also statistically different from the vehicle types. The rear-end crash risk model results indicated that the rear-end crash risk increases with the heavy vehicle percentage and the lane traffic flow rate. An interesting finding from the model results was that the lane closer to the work zone is strongly associated with the higher rear-end crash risk. Because of less braking capability, a truck has much higher risk of being involved in a rear-end crash than a car. The expressway work zone was further found to have significantly higher rear-end crash risk than the arterial work zone, which is consistent with the findings from the literature.
CHAPTER 8 WORK ZONE CASUALTY RISK ASSESSMENT

8.1 Introduction

Although a number of models have been developed for the crash risk assessment in previous studies, these models emphasize estimating the occurrence likelihood or frequency of work zone crash. However, in practice, traffic safety engineers seem to pay more attention on the casualty risk, e.g., the likelihood of a driver/passenger being killed or injured in a work zone, and on the relationship between the frequencies and consequences of work zone crashes. To date, few work zone risk assessment studies have been focused on the assessment of casualty risk by simultaneously taking into account the frequencies and consequences of work zone crashes.

Fortunately, the QRA model has been proved to be an effective methodology quantitatively to assess the overall safety level of hazards (NUREG, 1975). It allows a quantitative assessment of a facility’s risks, rendered by a broad range of accidents from frequent-minor to rare-major accidents. This chapter therefore aims to develop a probabilistic QRA model in an attempt to assess the casualty risk combining frequencies and consequences of work zone crashes. To achieve this objective, it is required to estimate the frequencies and consequences of all crash accident scenarios. Two risk expressions including individual risk and societal risk are used to express the casualty risk in work zone crashes. The individual risk is interpreted as the frequency
of a driver/passenger being killed or injured and the societal risk describes the relationship between the frequency and the total number of casualties.

8.2 Probabilistic Quantitative Risk Assessment Model Formulation

The occurrence of a work zone crash may lead to various consequences including fatalities or injuries. Hence, there would be a number of possible accident scenarios with distinct consequences for the work zone crash. The possible accident scenarios can be logically illustrated by an event tree diagram in which all possible paths following a top event (e.g. work zone crash) can be traced through intermediate events. The occurrence frequency of a particular accident scenario hence equals to the product of work zone crash frequency and the occurrence probability/likelihood of this scenario. The consequence of this accident scenario can be estimated by using the consequence estimation model. Therefore, the major task in the model formulation is to estimate the work zone crash frequency, to build an event tree and a consequence estimation model, and to calculate the casualty risk. A flowchart of the QRA model formulation is depicted in Figure 8.1, and its components are elaborated as follows.
8.2.1 Work Zone Crash Frequency

In general, work zone crash frequency increases with the work zone length \((L)\) and work zone duration \((D)\). In addition, traffic volume \((Q)\) and road type \((U)\) are another two factors influencing the work zone crash frequency. From a statistical viewpoint, the lower traffic volume on a work zone, the lesser effect on work zone crash it results in (Rouphail et al., 1988). Khattak et al. (2002) summarized that the work zone length, work zone duration, traffic volume and road type are important determinants of the crash frequency in work zone. Hence, the frequency of work zone crash is modeled as a function of these four determinants.

The functional form can be identified by using the least squares method based on the historical crash accident data shown in Table 8.1. To do so, the road type \((U)\) is first expressed by a binary variable, where \(U=1\) for an urban road and \(U=0\) for a rural road. The work zone crash frequency function can be formulated as the following
form:

$$\ln(f) = \alpha_0 + \sum_{i=1}^{3}(\alpha_i \times x_i) + \alpha_4 \times U + \varepsilon$$  \hspace{1cm} (8.1)$$

where $f$ is the work zone crash frequency, $\alpha_0, \alpha_i, i=1,..,4$ are the coefficients to be estimated, $\varepsilon$ is the random error term, and $x_i, i=1,2,3$ are the three explanatory variables associated with the work zone length, work zone duration and traffic volume, respectively.

<table>
<thead>
<tr>
<th>Length (mile)</th>
<th>Annual average daily traffic (veh/day)</th>
<th>Work zone duration (days)</th>
<th>Road Type</th>
<th>Work zone crash frequency (crash per work zone duration)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.12</td>
<td>113,108</td>
<td>304</td>
<td>1*</td>
<td>174</td>
</tr>
<tr>
<td>3.20</td>
<td>84,947</td>
<td>274</td>
<td>0</td>
<td>163</td>
</tr>
<tr>
<td>4.29</td>
<td>86,397</td>
<td>274</td>
<td>1</td>
<td>134</td>
</tr>
<tr>
<td>5.50</td>
<td>71,116</td>
<td>304</td>
<td>1</td>
<td>140</td>
</tr>
<tr>
<td>3.79</td>
<td>60,257</td>
<td>213</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>5.90</td>
<td>63,656</td>
<td>365</td>
<td>0</td>
<td>212</td>
</tr>
<tr>
<td>1.17</td>
<td>85,111</td>
<td>207</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>4.00</td>
<td>41,470</td>
<td>213</td>
<td>1</td>
<td>48</td>
</tr>
<tr>
<td>2.70</td>
<td>41,192</td>
<td>122</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>4.18</td>
<td>63,032</td>
<td>335</td>
<td>1</td>
<td>117</td>
</tr>
<tr>
<td>3.00</td>
<td>41,865</td>
<td>122</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>2.50</td>
<td>41,487</td>
<td>243</td>
<td>1</td>
<td>43</td>
</tr>
<tr>
<td>5.42</td>
<td>55,334</td>
<td>365</td>
<td>1</td>
<td>125</td>
</tr>
<tr>
<td>2.92</td>
<td>60,681</td>
<td>213</td>
<td>1</td>
<td>55</td>
</tr>
</tbody>
</table>

Note: * 1 for urban road; 0 for rural road

It is assumed that each of the three explanatory variables has three possible relations with respect to its relevant contributing factor: identical, logarithm and reciprocal. For example, the explanatory variable $x_1$ related to the work zone length can be expressed by $L$, $\ln(L)$ or $(1/L)$. There are $3 \times 3 \times 3 = 27$ combinations of forms for the explanatory variables shown in Eq. (8.1). The following trial-and-error method based on the multiple linear regression analysis is thus used to find the best
combination form. First, all possible combination forms are enumerated. For each combination form, the multiple linear regression method is applied for the available field data to estimate the coefficients $\alpha_0$, $\alpha_1$, $\alpha_2$, $\alpha_3$, and $\alpha_4$. According to the $p$-values of the estimated coefficients, a feasible combination form is defined as one with all $p$-values less than the significance level of 0.05. Second, the best combination form with the largest adjusted $R^2$ is chosen from all feasible combination forms.

<table>
<thead>
<tr>
<th>Combination form $$(x_1, x_2, x_3, x_4)$$</th>
<th>$p$-values for coefficients</th>
<th>Feasibility</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(L,D,Q,U)$</td>
<td>$&lt;0.01$ 0.34 $&lt;0.01$ $&lt;0.01$ 0.38</td>
<td>Feasible</td>
<td>88.9%</td>
</tr>
<tr>
<td>$(L,D,1/Q,U)$</td>
<td>$&lt;0.01$ 0.22 $&lt;0.01$ $&lt;0.01$ 0.46</td>
<td></td>
<td>92.9%</td>
</tr>
<tr>
<td>$(L,D,\ln(Q),U)$</td>
<td>$&lt;0.01$ 0.24 $&lt;0.01$ $&lt;0.01$ 0.42</td>
<td></td>
<td>91.7%</td>
</tr>
<tr>
<td>$(L,1/D,Q,U)$</td>
<td>$&lt;0.01$ 0.02 $&lt;0.01$ $&lt;0.01$ 0.04</td>
<td>Feasible</td>
<td>93.4%</td>
</tr>
<tr>
<td>$(L,1/D,1/Q,U)$</td>
<td>$&lt;0.01$ $&lt;0.01$ $&lt;0.01$ $&lt;0.01$ 0.09</td>
<td></td>
<td>94.7%</td>
</tr>
<tr>
<td>$(L,1/D,\ln(Q),U)$</td>
<td>0.02  $&lt;0.01$ $&lt;0.01$ $&lt;0.01$ 0.07</td>
<td></td>
<td>94.1%</td>
</tr>
<tr>
<td>$(L,\ln(D),Q,U)$</td>
<td>0.01  0.12 $&lt;0.01$ $&lt;0.01$ 0.15</td>
<td></td>
<td>92.1%</td>
</tr>
<tr>
<td>$(L,\ln(D),1/Q,U)$</td>
<td>0.71  0.06 $&lt;0.01$ $&lt;0.01$ 0.19</td>
<td></td>
<td>95.6%</td>
</tr>
<tr>
<td>$(L,\ln(D),\ln(Q),U)$</td>
<td>$&lt;0.01$ 0.07 $&lt;0.01$ $&lt;0.01$ 0.16</td>
<td></td>
<td>94.1%</td>
</tr>
<tr>
<td>$(\ln(L),D,Q,U)$</td>
<td>$&lt;0.01$ 0.40 $&lt;0.01$ $&lt;0.01$ 0.35</td>
<td></td>
<td>88.6%</td>
</tr>
<tr>
<td>$(\ln(L),D,1/Q,U)$</td>
<td>$&lt;0.01$ 0.22 $&lt;0.01$ $&lt;0.01$ 0.38</td>
<td></td>
<td>92.9%</td>
</tr>
<tr>
<td>$(\ln(L),D,\ln(Q),U)$</td>
<td>$&lt;0.01$ 0.27 $&lt;0.01$ $&lt;0.01$ 0.36</td>
<td></td>
<td>91.5%</td>
</tr>
<tr>
<td>$(\ln(L),1/D,Q,U)$</td>
<td>$&lt;0.01$ 0.03 $&lt;0.01$ $&lt;0.01$ 0.04</td>
<td>Feasible</td>
<td>92.7%</td>
</tr>
<tr>
<td>$(\ln(L),1/D,1/Q,U)$</td>
<td>$&lt;0.01$ $&lt;0.01$ $&lt;0.01$ $&lt;0.01$ 0.04</td>
<td>Feasible</td>
<td>95.4%</td>
</tr>
<tr>
<td>$(\ln(L),1/D,\ln(Q),U)$</td>
<td>0.03  0.01 $&lt;0.01$ $&lt;0.01$ 0.04</td>
<td>Feasible</td>
<td>94.5%</td>
</tr>
<tr>
<td>$(\ln(L),\ln(D),Q,U)$</td>
<td>$&lt;0.01$ 0.16 $&lt;0.01$ $&lt;0.01$ 0.13</td>
<td></td>
<td>91.7%</td>
</tr>
<tr>
<td>$(\ln(L),\ln(D),1/Q,U)$</td>
<td>0.52  0.07 $&lt;0.01$ $&lt;0.01$ 0.13</td>
<td></td>
<td>94.8%</td>
</tr>
<tr>
<td>$(\ln(L),\ln(D),\ln(Q),U)$</td>
<td>$&lt;0.01$ $&lt;0.01$ $&lt;0.01$ $&lt;0.01$ 0.13</td>
<td></td>
<td>93.8%</td>
</tr>
<tr>
<td>$(1/L,D,Q,U)$</td>
<td>$&lt;0.01$ 0.52 $&lt;0.01$ $&lt;0.01$ 0.37</td>
<td></td>
<td>88.2%</td>
</tr>
<tr>
<td>$(1/L,D,1/Q,U)$</td>
<td>$&lt;0.01$ 0.27 $&lt;0.01$ $&lt;0.01$ 0.38</td>
<td></td>
<td>92.6%</td>
</tr>
<tr>
<td>$(1/L,D,\ln(Q),U)$</td>
<td>$&lt;0.01$ 0.35 $&lt;0.01$ $&lt;0.01$ 0.37</td>
<td></td>
<td>91.1%</td>
</tr>
<tr>
<td>$(1/L,1/D,Q,U)$</td>
<td>$&lt;0.01$ 0.07 $&lt;0.01$ $&lt;0.01$ 0.06</td>
<td></td>
<td>91.3%</td>
</tr>
<tr>
<td>$(1/L,1/D,1/Q,U)$</td>
<td>$&lt;0.01$ 0.03 $&lt;0.01$ $&lt;0.01$ 0.06</td>
<td></td>
<td>94.4%</td>
</tr>
<tr>
<td>$(1/L,1/D,\ln(Q),U)$</td>
<td>0.08  0.04 $&lt;0.01$ $&lt;0.01$ 0.06</td>
<td></td>
<td>93.3%</td>
</tr>
<tr>
<td>$(1/L,\ln(D),Q,U)$</td>
<td>0.01  0.25 $&lt;0.01$ $&lt;0.01$ 0.14</td>
<td></td>
<td>91.0%</td>
</tr>
<tr>
<td>$(1/L,\ln(D),1/Q,U)$</td>
<td>0.62  0.11 $&lt;0.01$ $&lt;0.01$ 0.13</td>
<td></td>
<td>94.3%</td>
</tr>
<tr>
<td>$(1/L,\ln(D),\ln(Q),U)$</td>
<td>$&lt;0.01$ 0.14 $&lt;0.01$ $&lt;0.01$ 0.13</td>
<td></td>
<td>93.2%</td>
</tr>
</tbody>
</table>

Note: A combination of variable form is feasible only if all $p$-values are less than the level of significance 0.05.
The data shown in Table 8.1 are used to demonstrate the implementation of the trial-and-error method for determining the best functional form of Eq. (8.1). Table 8.2 gives all possible statistical analysis results obtained by the trial-and-error method. It shows that there are four feasible combination forms. Among these four feasible combination forms, the best one is \((\ln(L), 1/D, 1/Q, U)\) because it yields the largest adjusted \(R^2\) (95.4\%). Therefore, the crash frequency function can be calibrated by

\[
\ln(f) = \alpha_0 + \alpha_1 \ln(L) + \frac{\alpha_2}{D} + \frac{\alpha_3}{Q} + \alpha_4 \times U
\]  

(8.2)

where the estimations of coefficients \(\alpha_0, \alpha_1, \alpha_2, \alpha_3\) and \(\alpha_4\) and their statistical analysis are summarized as follows:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std Err</th>
<th>(p)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_0)</td>
<td>6.12</td>
<td>0.24</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>(\alpha_1)</td>
<td>0.429</td>
<td>0.13</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>-215</td>
<td>37.8</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>(\alpha_3)</td>
<td>-66468</td>
<td>11278</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>(\alpha_4)</td>
<td>-0.235</td>
<td>0.10</td>
<td>0.04</td>
</tr>
</tbody>
</table>

In other words, the work zone crash frequency function has the following form according to Eq. (8.2).

\[
f(L, D, Q, U) = e^{6.12 - 215/D - 66468/Q - 0.235U} \times L^{0.429}
\]  

(8.3)

It should be pointed out that the work zone crash frequency function shown by Eq. (8.3) is valid subject to the three conditions: \(1.17 \leq L \leq 5.9, 122 \leq D \leq 365\) and \(41,192 \leq Q \leq 113,108\) because the observed work zone length, work zone duration and traffic volume data are within these three ranges, respectively.
8.2.2 Event Tree Building & Occurrence Probability Calculation

In a work zone crash, the occurrence probability of an accident scenario will depend on the driver, vehicle, roadway, and environment conditions. An event tree is applied to determine all possible accident scenarios and their corresponding occurrence probabilities.

Since the occurrence likelihood of an accident scenario varies based on the age, crash unit, vehicle type, alcohol, light condition, crash type and crash severity, these seven contributing factors are regarded as intermediate events for the event tree. In general, the event tree is investigated from left to right. Thus, the event tree is started at the “Vehicle Crash” (top event) column and terminated at the “Severity” column, as shown in Figure 8.2. The event tree is decomposed into three sub-event trees, namely sub-event tree (a), sub-event tree (b) and sub-event tree (c). Sub-event tree (b) will continue from the intermediate event of vehicle type in the sub-event tree (a). Similarly, sub-event tree (c) will be linked to the event of light condition in the sub-event tree (b).
The table presents a work zone casualty risk assessment with the following columns:

- **Age (A)**: Categorized into two levels.
- **Crash Unit (CU)**: Categorized into five levels: 1-unit, >=2-unit, 3-unit, 2-unit, and 1-unit.
- **Vehicle Type (VT)**: Categorized into three levels: heavy vehicle, light vehicle, and heavy + light vehicles.

The diagram and sub-event trees illustrate the risk assessment process, with notes indicating the categorization of alcohol and light condition factors. The figure shows the structure of the event tree for work zone crashes.
types: young and others. This is because the specific safety needs of young (<25 year-old) might need to be separately addressed (Harb et al., 2008). Since heavy vehicles involved work zone crashes are more dangerous than non-heavy vehicle involved crashes (Pigman and Agent, 1990), the work zone crash is classified into two groups: i) heavy vehicle involved in crash and ii) non-heavy vehicle involved in crash. According to the number of vehicles involved, four types of crash units are listed in the “Crash Unit” column, which are 1-unit, 2-unit, 3-unit and >=4 unit, respectively. In addition, the light condition is categorized into daylight, dark with lighting and dark without light because the work zone crash is primarily occurred in these three types of light conditions. According to the crash severity, three types of work zone crashes comprising fatal crash, injury crash and property damage-only (PDO) crash are considered, as shown in Figure 8.2. In addition, all possible consequences of a work zone crash can be described as: i) deaths and injuries in a fatal crash; ii) injuries in an injury crash and iii) property damage only in a PDO crash, shown in the “Consequence Description” column.

A particular accident scenario depends on the path, which is represented by a specific sequence of intermediate events in the event tree. Therefore, the occurrence probability associated with a particular accident scenario can be simply determined by multiplying the probabilities of all intermediate events on that path. According to the event tree structure in Figure 8.2, there are a total of \(2 \times 4 \times 2 \times 2 \times 3 \times (2 + 1) = 288\) possible accident scenarios. Let \(p_{eqr}, \ r = 1, 2, \ldots, 288\), be the occurrence probability of the \(r^{th}\) accident scenario. Mathematically, it can be expressed by:
where \( E_k \) \((k=1,\ldots,7)\) is the intermediate event \( k \) along the path possessed by the \( r \)th accident scenario, \( P(E_k | E_{k-1}) \) is the probability that the event \( E_k \) is triggered by the event \( E_{k-1} \) along the path possessed by the \( r \)th accident scenario.

According to Eq.(8.4), it can be seen that the accuracy of estimated occurrence probability associated with the accident scenario depends on the probabilities of the intermediate events along the path. The probabilities of the intermediate events can be estimated from the historical data. However, the variation and uncertainty may exist in the probability estimation of some intermediate events. For example, the probability that a light vehicle collides with a fixed object, that is, the probability of “1-unit” event, can be estimated from the historical crash data. However, this probability may not be suitable to be regarded as a crisp value. This is because the value may vary due to the statistical variation in external conditions, such as the likelihood of a fixed object appearing in a work zone, or the variability the variable has. A second probability concept can be used to describe this variation/uncertainty as a probability distribution around a “point estimate” for the probability of the “1-unit” event. Therefore, the estimated probabilities of the intermediate events in the event tree are considered to be random variables, which can be represented by means of probability distribution in this study. The Monte Carlo sampling method is a well-recognized method addressing this problem.
8.2.3 Consequence Model

In general, the number of casualties depends on the number of crash units involved in a crash, vehicle type and crash severity. In addition, the speed and the emergency response time (ERT) also affect the consequence of a work zone crash. In this section, the consequence model is developed to evaluate the number of casualties for each accident scenario.

According to the nature of PDO crash, it is reasonably assumed that there are no casualties in a PDO crash. Regarding the estimation of casualties in a fatal or injury crash, these two kinds of severe crashes are separately analyzed. In a fatal work zone crash, passengers or drivers in the involved vehicles have the risk of both being killed and injured while only the risk of being injured is considered in an injury work zone crash. When the mean speed is $V_0$ km/h, the number of fatalities and injuries in a vehicle of type $i$, respectively denoted by $N_{Fi}$ and $N_{Ii}$, can be calculated by

$$N_{Fi} = \begin{cases} N_{FFi}, & \text{in a fatal crash accident} \\ 0, & \text{in an injury crash accident} \end{cases}$$ \hfill (8.5)

$$N_{Ii} = \begin{cases} N_{IFi}, & \text{in a fatal crash accident} \\ N_{IIi}, & \text{in an injury crash accident} \end{cases}$$ \hfill (8.6)

$$i = \begin{cases} 1, & \text{for the light vehicle} \\ 2, & \text{for the heavy vehicle} \end{cases}$$ \hfill (8.7)

where $N_{FFi}$ = the number of fatalities in a vehicle of type $i$ in a fatal work zone crash;

$N_{IFi}$ = the number of injuries in a vehicle of type $i$ in a fatal work zone crash;

$N_{IIi}$ = the number of injuries in a vehicle of type $i$ in an injury work zone crash.

Further, the number of fatalities and injuries should fulfill the following threshold limits:

In a fatal crash
Chapter 8 Work Zone Casualty Risk Assessment

\[ N_i = N_{F_i} + N_{H_i} \quad (8.8) \]

\[ 0 < N_{F_i} \leq N_i \quad (8.9) \]

**In an injury crash**

\[ 0 < N_{H_i} \leq N_i \quad (8.10) \]

where the parameter \( N_i \) is the average occupancy of a vehicle of type \( i \). Constraints (8.8) and (8.9) imply that the passengers or drivers have the risk of being killed in a fatal work zone crash. Constraint (8.10) ensures that the number of injuries should not exceed the vehicle occupancy.

**Effects of speed and ERT**

The speed can affect the number of casualties in a work zone crash. Intuitively, the greater the speed in a crash accident, the more casualties it will lead to. The Power Model, originally proposed by Nilsson (1981, 2004), first addresses the following quantitative relationship between the number of casualties and speed.

\[ \frac{n_i}{n_0} = \left( \frac{V_i}{V_0} \right)^\alpha \quad (8.11) \]

where \( n_i \) is the number of causalities per involved vehicle at the mean speed of \( V_i \); \( n_0 \) is the number of causalities per involved vehicle at the mean speed of \( V_0 \); the exponent parameter is \( \alpha = 4 \) in fatal accidents and \( \alpha = 2 \) in injury accidents.

Elvik et al. (2004) demonstrated that the Power Model is in a very good agreement with the empirical experience for injury accidents and fatal accidents. To test the relative plausibility of the Power Model, other alternative models including a linear model and a logistic model are examined by Elvik et al. (2004). They pointed out that the linear model is highly implausible and the logistic model with exponent depending on initial speed does not perform better than the Power Model because the gain in precision is too small to be justified. In addition, European Commission (1999)
explicitly pointed out that the relationship between speed and accident severity can be expressed by the Power Model. However, the Power Model with the exponents of four and two may not be the best. Elvik et al. (2004) assembled a large dataset from 97 published studies containing 460 results to reformulate the Power Model. They presented the best estimates of exponents and the corresponding confidence interval for the causalities in the fatal and injury accidents. Because of its plausibility and simplicity, the Power Model has been found a widespread use to quantify the effects of speed on road safety (Cameron and Elvik, 2008; Evans, 1991; Kallberg, 1998; Woolley, 2005).

In this chapter, the Power Model is also employed to estimate the effect of speed on the consequence of an accident scenario. Suppose that the mean speed has the same effects on the risk of being killed or injured for a driver/passenger in different vehicle types. Therefore, it can be obtained:

\[
N_{i}^{(1)} = \left( \frac{V_1}{V_0} \right)^{\alpha_i} \times N_{F_i}
\]  
(8.12)

where \( N_{i}^{(1)} \) is the number of fatalities in an involved vehicle at the mean speed of \( V_1 \), \( N_{F_i} \) is the number of fatalities in a vehicle of type \( i \) at the mean speed of \( V_0 \), \( \alpha_i \) is the best-estimated exponents in fatal crash.

Taking into account the threshold limit shown by Eq. (8.9), the number of fatalities in Eq. (8.12) should be

\[
N_{i}^{(1)} = \min \left\{ \left( \frac{V_1}{V_0} \right)^{\alpha_i} \times N_{F_i}, N_i \right\}
\]  
(8.13)

where \( N_i \) is the average occupancy of a vehicle of type \( i \).

The ERT is crucial for the timely delivery of emergency medical services
(EMS) to accident casualties. It is expected to have an impact on the number of fatalities because little or no first aid to serious injurers may lead to more deaths in the fatal crashes. Evanco (1996) addressed a linear relationship between the ERT and the number of fatalities. After taking the effect of ERT into consideration, the number of fatalities in an involved vehicle can be evaluated by the following linear function:

\[ N_{ri}^{(2)} = \left( \beta + (1 - \beta) \times \frac{T_1}{T_0} \right) \times N_{ri}^{(1)} \]  

(8.14)

where \( N_{ri}^{(2)} \) represents the number of fatalities taking into account the impact of ERT; \( T_0 \) is the normal mean response time (min); \( T_1 \) is the actual mean response time (min), \( \beta \) is the parameter to be estimated; \( N_{ri}^{(1)} \) is the number of fatalities in an involved vehicle of type \( i \).

Also, considering the above threshold limit, the number of fatalities in Eq. (8.14) should satisfy

\[ N_{ri}^{(2)} = \min \left\{ \left( \beta + (1 - \beta) \times \frac{T_1}{T_0} \right) \times N_{ri}^{(1)}, N_i \right\} \]  

(8.15)

where \( N_{ri}^{(1)} \) is the number of fatalities in an involved vehicle; \( N_i \) is the average occupancy of a vehicle of type \( i \).

An assumption is made that the ERT has no impact on the number of injuries in an injury crash accident. Hence, the number of injuries after taking into account the effects of speed and ERT can be calculated by:

\[ N_{ri}^{(2)} = \begin{cases} N_i - N_{ri}^{(2)} & \text{, in a fatality crash accident} \\ \min \left\{ \left( \frac{V}{V_0} \right)^{a_2} \times N_{ri}, N_i \right\} & \text{, in an injury crash accident} \end{cases} \]  

(8.16)
where \( N^{(2)}_{i} \) is the number of injuries in a vehicle of type \( i \) after taking into account the effects of speed and \( ERT \); \( N_{i} \) is the number of injuries in a vehicle of type \( i \); \( N_{i} \) is the average occupancy of a vehicle of type \( i \); \( \alpha_{2} \) is the best-estimated exponents in an injury crash accident.

The total number of fatalities and injuries involved in a multi-unit vehicle crash can be obtained using the simple method of algebraic sum. Let \( N_{TF} \) and \( N_{FI} \) respectively represent the total number of fatalities and injuries in the \( r \)th accident scenario, \( r = 1, 2, \ldots, 288 \), and they could mathematically be calculated by:

\[
N_{TF} = m_{1} \cdot N^{(2)}_{F_{1}} + m_{2} \cdot N^{(2)}_{F_{2}} \quad (8.17)
\]

\[
N_{FI} = m_{1} \cdot N^{(2)}_{I_{1}} + m_{2} \cdot N^{(2)}_{I_{2}} \quad (8.18)
\]

where \( N_{TF} \) and \( N_{FI} \) are respectively the total number of fatalities and injuries in the \( r \)th accident scenario; \( N^{(2)}_{F_{1}} \) and \( N^{(2)}_{F_{2}} \) are respectively the number of fatalities and injuries of a light vehicle; \( N^{(2)}_{I_{1}} \) and \( N^{(2)}_{I_{2}} \) are respectively the number of fatalities and injuries of a heavy vehicle; \( m_{1} \) and \( m_{2} \) represent the number of light vehicles and heavy vehicles involved in the \( r \)th accident scenario, respectively.

### 8.2.4 Quantitative Casualty Risk Expressions

It is now widely recognized that the risk from an accident refers to a function of the likelihood of occurrence of possible undesired events and the magnitude of their associated consequences (Borysiewicz et al., 2006). This chapter also adopts two commonly used ways to quantitatively express the casualty risk caused by work zone crashes. One is the individual risk and the other is the societal risk. The expressions of
individual and societal risks incorporate the frequencies and consequences of all the accident scenarios that have been identified.

**Individual risk**

Considine (1984) defines the individual risk as the likelihood or frequency of the fatality/injury occurring to a person in the vicinity of a hazard. In this study, individual risk is defined as the frequency of a passenger or a driver being killed or injured when traveling through work zone. It includes two aspects: (i) individual fatality risk and (ii) individual injury risk. The individual fatality and injury risks from all accident scenarios, respectively denoted by $IR_f$ and $IR_i$, can be calculated by:

\[
IR_f = \frac{f_{\text{crash}}(Q,D,L,U) \cdot \sum_{r=1}^{288} (p_{\text{seq}}N_{TR})}{Q((1-hv)N_i + hv \cdot N_z)}
\]

(8.19)

\[
IR_i = \frac{f_{\text{crash}}(Q,D,L,U) \cdot \sum_{r=1}^{288} (p_{\text{seq}}N_{TI})}{Q((1-hv)N_i + hv \cdot N_z)}
\]

(8.20)

where $IR_f$ and $IR_i$ are the individual fatality and injury risks from all accident scenarios, respectively; $f_{\text{crash}}(Q,D,L,U)$ is the work zone crash frequency; $p_{\text{seq}}$ is the occurrence probability of the $r^{th}$ accident scenario; $N_{TR}$ and $N_{TI}$ are respectively the total number of fatalities and injuries in the $r^{th}$ accident scenario; $N_i$ and $N_z$ are respectively the average number of persons in a light vehicle and a heavy vehicle; $hv$ is the heavy vehicle percentage; $Q((1-hv)N_i + hv \cdot N_z)$ represents the total number of vehicle occupants traveling across work zone.

**Societal risk**

The individual risk implies that the risk can be aggregated into a single
number. Obviously, a single number cannot address the relationship between the frequency and the total number of casualties caused by a work zone crash. Therefore, the societal risk is employed to complement the individual risk measure in this study. It is a measure of risk to a group of people.

The most common form of presentation of societal risk is the $F/N$ curve, which illustrates the relationship between the crash consequence and the corresponding occurrence frequency. The number of casualties $x$ (e.g., fatalities and injuries) and the corresponding frequencies $F_N(x)$ is shown on the abscissa and ordinate, respectively. Mathematically, $F/N$ curve can be expressed by

$$F_N(x) = P(x \geq N)$$  \hspace{1cm} (8.21)

where $x$ is the number of casualties caused by a work zone crash, and $N$ is a given value; $P(x \geq N)$ can be described as the frequency that the number of casualties caused by work zone crashes, $x$, is not less than the number $N$.

### 8.3 Numerical Example

One numerical example is carried out to test the proposed QRA model for assessing the casualty risk caused by vehicle crashes in a work zone. Considering a work zone with the total length of 4.1 kilometers (e.g., $L \approx 2.6$ mile) and the work zone duration of 130 days (e.g., $D=130$ d). The work zone is located at an urban road of the Southeast Michigan State.

Suppose that the work zone crash frequency satisfies the relationship shown by Eq.(8.3), the frequency of work zone crash is 23.66 crashes per work zone duration (130 d) if the average daily traffic volume $Q$ is 45,000 vehicles per day in this example.
8.3.1 Data

The Southeast Michigan Traffic Crash Records Database for the years from 1999 to 2008 is utilized and is originally obtained from the Southeast Michigan Council of Governments. The database contains 89 fatal work zone crashes involved fatalities and injuries, 10,142 injury work zone crashes involved injuries and 35,036 PDO work zone crashes. From this 10-year’s work zone crash database, it can be found that the probabilities that more than one heavy vehicle are respectively involved in 2-Unit, 3-Unit, and >=4-Unit work zone crashes are extremely low. Therefore, only one heavy vehicle is considered in a heavy vehicle-involved crash. In addition, it does not take into account the crashes involving more than 4-unit vehicles because such crashes occupy extremely low percentage (less than 1.0E-04) among all the recorded work zone crashes. Therefore, the number of light vehicles $m_1$ and the number of heavy vehicles $m_2$ in each accident scenario are shown in Figure 8.2.

Assuming that the values of $V_0$ and $T_0$ are respectively 60 km/h and 5.2 min, the average number of fatalities and injuries per vehicle in a fatal/injury crash accident can be estimated using the linear regression method from the database, which are shown in Table 8.3. To apply the proposed model, this chapter takes a set of values of the parameters $V_1$, $T_1$, $\alpha_1$, $\alpha_2$, $\beta$, $hv$ (see in Table 8.3). Car and motorcycle are considered as the major two types of light vehicle and the heavy vehicle consists of van and truck. The average value of car and motorcycle occupancy is thus regarded as the light vehicle occupancy ($N_1$). The heavy vehicle occupancy ($N_2$) approximates the average occupancy of van and truck.
Table 8.3 Input parameters for the numerical example

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Descriptions</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$hv$</td>
<td>Percentage of heavy vehicles</td>
<td>17.0%</td>
</tr>
<tr>
<td>$N_1$</td>
<td>Average number of persons in a light vehicle</td>
<td>1.54</td>
</tr>
<tr>
<td>$N_2$</td>
<td>Average number of persons in a heavy vehicle</td>
<td>1.70</td>
</tr>
<tr>
<td>$N_{FF_1}$</td>
<td>Expected number of fatalities of a light vehicle involving a fatal crash accident</td>
<td>0.54</td>
</tr>
<tr>
<td>$N_{FF_2}$</td>
<td>Expected number of fatalities of a heavy vehicle involving a injury crash accident</td>
<td>0.62</td>
</tr>
<tr>
<td>$N_{II_1}$</td>
<td>Expected number of injuries of a light vehicle involving an injury crash accident</td>
<td>0.65</td>
</tr>
<tr>
<td>$N_{II_2}$</td>
<td>Expected number of injuries of a heavy vehicle involving an injury crash accident</td>
<td>0.90</td>
</tr>
<tr>
<td>$V_1$</td>
<td>Actual mean speed</td>
<td>65km/h</td>
</tr>
<tr>
<td>$T_1$</td>
<td>Actual mean value of EMS response time</td>
<td>4.8 min</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>The best estimate exponents in fatal crash</td>
<td>4.5$^a$</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>The best estimate exponents in injury crash</td>
<td>2.7$^a$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>The parameter in Eq. (8.14)</td>
<td>0.73$^b$</td>
</tr>
</tbody>
</table>

Note: $^a$ source from Elvik et al. (2004), confidence interval: $\alpha_1$ (4.1–4.9); $\alpha_2$ (0.9–4.5); $^b$ source from Evanco (1996)

Occurrence probabilities of intermediate events in the event tree

Table 8.4 reports the means and relative standard deviations of the probabilities of all intermediate events from the database. It can be seen that 2-unit crashes are the primary crashes among all work zone crashes because of high percentage of 0.7404 (for age less than 25) and 0.7351 (for age larger than 25). The relative standard deviations of the probabilities of intermediate events of vehicle type ($VT$), crash type ($CT$) and severity ($S$) are large. It implies that the estimated probabilities of these intermediate events have large variations (uncertainties). Therefore, the estimated probabilities of intermediate events $VT$, $CT$ and $S$ are described by probabilistic distributions in the context of event tree analysis. On the
other hand, the occurrence probabilities of intermediate events of age \((A)\), crash unit \((CU)\), alcohol \((AL)\) and light condition \((LC)\) that have small relative standard deviations are assigned by sample means.

<table>
<thead>
<tr>
<th>Event</th>
<th>Acronym</th>
<th>Probability of occurrence</th>
<th>pdf</th>
<th>Mean</th>
<th>Relative standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>(A)</td>
<td>(p(A))</td>
<td>—</td>
<td>6.795E-01</td>
<td>1.10E-02</td>
</tr>
<tr>
<td>Crash Unit</td>
<td>(CU_1)</td>
<td>(p(CU_1))</td>
<td>—</td>
<td>1.062E-01</td>
<td>1.02E-02</td>
</tr>
<tr>
<td></td>
<td>(CU_2)</td>
<td>(p(CU_2))</td>
<td>—</td>
<td>7.404E-01</td>
<td>1.74E-03</td>
</tr>
<tr>
<td></td>
<td>(CU_3)</td>
<td>(p(CU_3))</td>
<td>—</td>
<td>1.217E-01</td>
<td>9.78E-03</td>
</tr>
<tr>
<td></td>
<td>(CU_4)</td>
<td>(p(CU_4))</td>
<td>—</td>
<td>3.170E-02</td>
<td>1.81E-02</td>
</tr>
<tr>
<td></td>
<td>(CU_5)</td>
<td>(p(CU_5))</td>
<td>—</td>
<td>1.791E-01</td>
<td>1.74E-02</td>
</tr>
<tr>
<td></td>
<td>(CU_6)</td>
<td>(p(CU_6))</td>
<td>—</td>
<td>7.351E-01</td>
<td>2.22E-02</td>
</tr>
<tr>
<td></td>
<td>(CU_7)</td>
<td>(p(CU_7))</td>
<td>—</td>
<td>6.950E-02</td>
<td>8.79E-03</td>
</tr>
<tr>
<td></td>
<td>(CU_8)</td>
<td>(p(CU_8))</td>
<td>—</td>
<td>1.630E-02</td>
<td>1.39E-02</td>
</tr>
<tr>
<td>Vehicle Type</td>
<td>(VT_1)</td>
<td>(v_1) (p(v_1))</td>
<td>—</td>
<td>9.844E-01</td>
<td>5.29E-02</td>
</tr>
<tr>
<td></td>
<td>(VT_2)</td>
<td>(v_2) (p(v_2))</td>
<td>—</td>
<td>9.335E-01</td>
<td>6.21E-02</td>
</tr>
<tr>
<td></td>
<td>(VT_3)</td>
<td>(v_3) (p(v_3))</td>
<td>—</td>
<td>9.419E-01</td>
<td>6.05E-02</td>
</tr>
<tr>
<td></td>
<td>(VT_4)</td>
<td>(v_4) (p(v_4))</td>
<td>—</td>
<td>9.220E-01</td>
<td>8.86E-02</td>
</tr>
<tr>
<td></td>
<td>(VT_5)</td>
<td>(v_5) (p(v_5))</td>
<td>—</td>
<td>9.279E-01</td>
<td>7.73E-02</td>
</tr>
<tr>
<td></td>
<td>(VT_6)</td>
<td>(v_6) (p(v_6))</td>
<td>—</td>
<td>8.281E-01</td>
<td>8.86E-02</td>
</tr>
<tr>
<td></td>
<td>(VT_7)</td>
<td>(v_7) (p(v_7))</td>
<td>—</td>
<td>8.968E-01</td>
<td>7.04E-02</td>
</tr>
<tr>
<td></td>
<td>(VT_8)</td>
<td>(v_8) (p(v_8))</td>
<td>—</td>
<td>8.475E-01</td>
<td>6.89E-02</td>
</tr>
<tr>
<td>Alcohol</td>
<td>(AL)</td>
<td>(p(AL))</td>
<td>—</td>
<td>9.632E-01</td>
<td>3.54E-03</td>
</tr>
<tr>
<td>Light Condition</td>
<td>(LC_1)</td>
<td>(p(LC_1))</td>
<td>—</td>
<td>7.653E-01</td>
<td>1.84E-02</td>
</tr>
<tr>
<td></td>
<td>(LC_2)</td>
<td>(p(LC_2))</td>
<td>—</td>
<td>1.655E-01</td>
<td>6.78E-03</td>
</tr>
<tr>
<td>Crash Type</td>
<td>(CT)</td>
<td>(v_9) (p(v_9))</td>
<td>—</td>
<td>7.869E-01</td>
<td>1.37E-01</td>
</tr>
<tr>
<td>Severity</td>
<td>(S)</td>
<td>(v_{10}) (p(v_{10}))</td>
<td>—</td>
<td>1.179E-02</td>
<td>2.92E-01</td>
</tr>
</tbody>
</table>

The @Risk software is allowed to fit probability distributions to the sample
data. This software makes use of the chi-squared fit statistic to measure how good the distribution fits the sample data. The smaller the value of chi-squared statistic is, the better the fit.

Figure 8.3 presents the best-fitted probability density functions of occurrence probabilities of intermediate events—vehicle type (VT), crash type (CT) and severity (S). In Figure 8.3, \( p(v_j), j = 1, \ldots, 8 \) represent probability density functions for the eight states of the event vehicle type (VT). \( p(v_9) \) and \( p(v_{10}) \) are respectively probability density functions in the events of crash type (CT) and severity (S). After the probability density functions \( p(v_j), j = 1, \ldots, 10 \) are determined, the propagation of the uncertainty of intermediate events VT, CT and S can be performed by resorting to the Monte Carlo sampling method. In the \( i^{th} \) Monte Carlo simulation, \( i = 1, 2, \ldots, 10000 \), the occurrence probabilities \( (v_1^i, \ldots, v_{10}^i) \) of the event vector \( (VT_1, \ldots, VT_8, CT, S) \) can be sampled from the corresponding probability density functions.
Figure 8.3 Probability density functions for the probabilities of intermediate events in the event tree

Chapter 8 Work Zone Casualty Risk Assessment
8.3.2 Results & Discussions

Given all the parameters required by the proposed model, the individual and societal risks that combine the frequencies and consequences of all accident scenarios could be calculated. For each realization of Monte Carlo simulation, the proposed model can yield one combination result, comprising the number of casualties and frequency, associated with a particular scenario. Percentile ranks are used to clarify the interpretation of all possible results. Figure 8.4 reports the percentile-based individual fatality and injury risks that a driver or passenger might experience during the work zone period. In Figure 8.4 (a), the 25th, 50th and 75th percentile of individual fatality risk are respectively $1.07 \times 10^{-6}/D, 1.33 \times 10^{-6}/D$ and $1.58 \times 10^{-6}/D$. Here, $D$ is the work zone duration, which is equal to 130 days. It implies that a passenger has the likelihood of $1.58 \times 10^{-6}$ of being killed when s/he is traveling through the work zone at a 75% confidence level. By comparing Figure 8.4 (a) with Figure 8.4 (b), it is shown that the passenger has a higher injury risk than fatality risk. The result is consistent with the fact that injuries are a relatively high percentage of the overall traffic accident casualties.

![Figure 8.4 Analysis of individual risk in work zone](image-url)
Figure 8.5 (a) and (b) depicts the relationship between the frequency and the number of casualties in a work zone crash. It can be clearly seen that the frequency decreases with the number of fatalities/injuries resulted from a work zone crash. It implies that the major crash accident (e.g. large number of fatalities or injuries) has a relatively low occurrence frequency than the minor accident in the work zone.

**Figure 8.5 Analysis of F/N curve for casualties in work zone**

From the practical point of view, the selection of $F/N$ curve should be made based on the worst case. This is because the occurrence probabilities of intermediate events in the event tree have variations/uncertainties, which might lead to underestimated results. In view of highly safety requirement, the conservative attitude is recommended in assessing the work zone risk. Therefore, the 95th percentile of $F/N$ curve could be considered as the societal risk caused by work zone crashes (see in Figure 8.5 (a) and (b)).
Impact of uncertainty associated with the probability of intermediate event

The larger uncertainties associated with the occurrence probabilities of intermediate events may cause higher degree of uncertainty associated with the output (e.g., individual fatality/injury risk). To evaluate the impact of the uncertainty with respect to the specified event (e.g., $S$), the corresponding occurrence probability is still represented by means of the probability distribution. Simultaneously, the occurrence probabilities of other intermediate events (e.g., $VT$ and $CT$) with large variations are assigned by the sample mean values.

![Diagram](image)

**Figure 8.6** Impact of uncertainty associated with the probability of intermediate event

Figure 8.6 (a) and (b) respectively describes the impact of uncertainty of event occurrence on the degree of uncertainty associated with the individual fatality and
injury risks. Intuitively, the uncertainty associated with the probability of intermediate event $S$ can cause a significant degree of uncertainty associated with the individual fatality risk. This is because the occurrence frequency of a driver being killed is greatly affected by the probability of intermediate event $S$. Similarly, the uncertainty of individual injury risk is significantly affected by the event $CT$.

In order quantitatively to determine the degree of uncertainty associated with the output of individual risk, the ratio of upper bound to lower bound is introduced in this section. Let $\rho(e)$ denote the uncertainty ratio of upper bound to lower bound for the uncertainty of outcome caused by the uncertain event $e$, and it can then be expressed by:

$$\rho(e) = \frac{X_{0.95}(e)}{X_{0.05}(e)}$$

(8.22)

where $X_{0.95}(e)=$ upper bound based on 95% percentile for event $e$; $X_{0.05}(e)=$ lower bound based on 5% percentile for event $e$.

The larger the uncertainty ratio is, the bigger uncertainty of the outcome will be resulted from the uncertainty event $e$.

Figure 8.6 (a) shows that the intermediate event $S$ leads to the highest degree of uncertainty associated with the individual fatality risk because the corresponding uncertainty ratio $\rho(S)$ is the largest, up to 2.88. Instead of the intermediate event $S$, the intermediate event $CT$ has the biggest impact on the degree of uncertainty associated with the individual injury risk, shown in Figure 8.6 (b). The degree of uncertainty associated with the individual fatality/injury risk is slightly affected by the intermediate event $VT$. The major reason is that the relative standard deviation of the estimated value for the probability of the intermediate event $VT$ is relatively small, as compared with the intermediate events $CT$ and $S$. 
Impact of *ERT* and speed on individual risk

As mentioned above, both the *ERT* and speed have impacts on the crash consequence in work zone.

![Figure 8.7 Influence of ERT on individual risk](image)

In Figure 8.7 (a), it can be seen that the individual fatality risk decreases with the *ERT*, whereas Figure 8.7 (b) shows that the individual injury risk is insignificantly affected by the *ERT*. This is because a delayed *ERT* decreases the chance of survival of passengers or drivers who are seriously injured in a fatal crash. It can be concluded that more deaths would be caused in work zone crashes if the local emergency medical service were low-efficiency.
Figure 8.8 Influence of speed on individual risk

Figure 8.8 demonstrates that the traveling speed poses a great influence on the individual fatality risk as well as the injury risk. There will be a 62% decrease of individual fatality risk and a 44% reduction of individual injury risk if the speed is slowed down by 20%. However, there will only be a 5% reduction of individual fatality risk and a 0.05% reduction of individual injury risk if the ERT is also reduced by 20%. Therefore, it can be concluded that reducing traveling speed is a more effective way to reduce the driver/passenger’s fatality and injury risks.

8.4 Summary

A novel probabilistic QRA model has been proposed to assess the casualty risk that combines frequency and consequence of work zone crash in this chapter. Since a number of accident scenarios may be possible in a work zone crash, the model
employs an event tree diagram to identify all possible accident scenarios. Seven intermediate events comprising age ($A$), crash unit ($CU$), vehicle type ($VT$), alcohol ($AL$), light condition ($LC$), crash type ($CT$) and severity ($S$) have been considered in the event tree analysis. The accuracy of the estimation of occurrence likelihood associated with a particular accident scenario is dependent upon the probability of each intermediate event, whereas the estimated probability values may have large variations. Therefore, the estimated values of probabilities for the intermediate events have been formulated as random variables. In contrast to simply categorizing consequence in previous studies, the probabilistic QRA model has applied a consequence model to estimate the quantitative consequence in terms of the number of casualties in each accident scenario. Two risk expressions have been utilized to express the casualty risk.

A numerical example was carried out and the Southeast Michigan work zone crash data was utilized to calibrate the proposed model. The estimated probability of intermediate events $VT$, $CT$ and $S$ with large variations are represented by means of probability distribution. The numerical example showed that the proposed probabilistic QRA model has the capability of reporting individual risk and societal risk in work zone. It also presents that the occurrence probability of the intermediate event $S$ with a large statistical variation causes a big uncertainty on the individual fatality risk. Slowing down speed is found to be more effective than reducing the $ERT$ in mitigating the individual fatality and injury risks.
CHAPTER 9 CONCLUSIONS

9.1 Outcomes and Contributions

Various work zone projects, such as pothole patching, roadside tree trimming and repairing damage to the roads, are necessary to maintain a good level of service for a road system. However, the presence of work zones can lead to high costs, increase severe crash risk and result in significant traffic delays. It is hence necessary to estimate traffic delay, analyze the optimal operational strategy and evaluate work zone risk before taking effective measures to lower work zone costs and reduce the negative traffic impacts.

However, it can be seen from the literature review that there are many limitations in the current models for each of the three main issues—traffic delay estimation, operational strategy analysis and risk assessment. For example, previous CA models assumed a constant randomization probability and an unrealistic lane change duration so that they cannot provide the realistic traffic behavior. The purpose of this research was therefore to propose new models and methodologies for the three issues. The contributions of this research can be summarized as follows:

- It could be very useful in helping land transport authorities to accurately estimate traffic delay at work zones.
- The proposed decision tree-based model could be a good alternative for traffic engineers to estimate work zone capacity because of its high estimation accuracy and ease of use.
• Contractors will be able quickly to determine the best operational strategy by using the proposed work zone cost minimization models.

• A new approach has been proposed for assessing work zone rear-end crash risk using work zone traffic data, when the historical accident data is unavailable.

• This research also makes an initiative for the assessment of casualty risk by simultaneously taking into account the frequencies and consequences of work zone crashes.

9.1.1 Work Zone Traffic Delay Estimation

Because of its simplicity and high degree of accuracy, the Cellular Automata (CA) model is considered to be a promising model for microscopic traffic simulation. Therefore, a heterogeneous Cellular Automata (HCA) model was developed to accurately estimate traffic delay at work zones. In the HCA model, the forwarding rules and lane changing rules are used to update the longitudinal and lateral vehicular behavior, respectively. Due to the unique traffic flow characteristics at work zones, the following modifications were made:

• Since drivers’ acceleration-deceleration behavior varies with work zone configuration and traffic flow, the randomization parameter is no longer a constant value in the HCA model. It is formulated as a function of the activity area length, the transition area length and the volumes of different types of vehicles traveling through work zone.
• The existing forwarding rules are inadequate to calculate the available front gap of the front vehicle in the blocked lane. Hence, a supplementary forwarding rule was added for the front vehicle in the blocked lane to calculate its available front gap.

• A realistic lateral speed and position updating rule was added so that the simulated lane change duration is close to reality.

• Compared with normal road traffic, work zone forbids lane changing from the through lane to the closed lane in the transition area. However, the existing lane changing rules are unable to cope with this realistic constraint. Therefore, two new lane change constraint rules have been added.

After the model was formulated, it was then calibrated and validated microscopically and macroscopically using real Singapore work zone data. Both the microscopic and macroscopic validation results show that the HCA model simulates the heterogeneous traffic in work zone well. After the validation, the model was applied to estimate traffic delay at work zones. The estimated delays were consistent with the field data to a high degree of accuracy. For the purpose of model performance comparison, the microscopic simulation software PARAMICS was also applied to estimate traffic delay. The HCA model was found to perform better in terms of accuracy.

Finally, an impact analysis was carried out to examine the marginal effects of the activity area length, transition area length, traffic flow and heavy vehicle percentage by using the HCA model. The results show that the transition area length
has a larger impact than the activity area length on traffic delay, especially in light traffic conditions.

### 9.1.2 Optimal Subwork Zone Operational Strategy

The deterministic queuing model is widely employed to determine the optimal subwork zone operational strategy because of its simplicity. Work zone capacity is a key parameter of the deterministic queuing model and the ability to estimate it accurately is thus imperative for traffic engineers. However, previous work zone capacity models and guidelines do not provide accurate estimates because they neglect to include important influencing factors and ignore the variable interaction problem. Therefore, this research has developed a decision tree-based model to estimate work zone capacity taking all important factors into consideration. To find an optimal decision tree, the $F$-test splitting criterion was employed to split nodes to grow the tree and a post-pruning approach is employed to prune the grown decision tree. Work zone capacity data collected from fourteen states and cities are used to train, check and evaluate the model.

The statistical comparison results demonstrate that the decision tree-based model outperforms the existing work zone capacity models and guidelines in terms of work zone capacity estimation accuracy. Compared with the existing models and guidelines, the new model also has high ease of use because traffic engineers can easily estimate the capacity of a given work zone by tracing a path down the tree to a
terminal node. Because of this high estimation accuracy and ease of use, the decision
tree-based model is then applied to determine an optimal subwork zone operational
strategy.

In order to determine an optimal subwork zone operational strategy, a
non-differentiable minimization model, with the objective of minimizing the total
work zone cost, was thus presented. In this model, the deterministic queuing model
and the HCA model were both applied to estimate total user delay. This total work
zone cost minimization model has remedied a few of the flaws in the queuing delay
and moving delay estimation formulae used in the existing subwork zone operational
strategy models. This model is also the first to take into account variable traffic speeds
and a uniform subwork zone length constraint to determine an optimal subwork zone
strategy. For comparison, this thesis also proposed a variation of the minimization
model without the uniform length constraint. It was found that the optimal operational
strategy without the uniform length constraint has a lower total work zone cost.

The total work zone cost can be regarded as the system cost, integrating the
costs of two different parties—the work zone contractor and the road user. From the
systemic viewpoint, there is a need to minimize the total work zone cost. However,
different parties have different objectives. With the objective of profit maximization,
work zone contractors want to minimize the total work zone maintenance cost, while
road users are only concerned about their travel delay in work zone. These two
objectives contradict each other to some extent. For the benefits of road users, land
transport authorities impose queue length and travel delay constraints on work zone
projects. Therefore, this thesis also proposed a total maintenance cost minimization model, subject to these two constraints, which determines the optimal subwork zone operational strategy from the contractor’s standpoint. The results show that the two constraints have a significant impact on the optimal solution. In addition, the optimal subwork zone operational strategy from the contractor’s standpoint is not always the same as that from the systemic viewpoint.

9.1.3 Risk Assessment at Work Zones

Most of the previous crash risk assessment studies rely extensively on historical accident data. However, poor quality historical accident data could result in biased results and these models are unable to assess the crash risk at newly-proposed work zones for which no historical data are available. Therefore, this research presents a new approach to assessing crash risks at work zones. Other types of data, such as work zone traffic data, are exploited to evaluate work zone crash risk. The deceleration rate to avoid the crash ($DRAC$) is used in measuring the rear-end crash risk. Based on arterial and expressway work zone traffic data in Singapore, three rear-end crash risk models were developed to examine the relationship between rear-end crash risk and its contributing factors: i) lane traffic flow rate; ii) lane traffic speed; iii) heavy vehicle percentage, iv) lane position and v) vehicle type.

From the one-way ANOVA test results, it can be seen that the rear-end crash risk estimated based on work zone traffic data is statistically different from the lane
positions at the expressway and arterial work zones. The rear-end crash risk at the arterial work zone is also statistically different from the vehicle types. As expected, the results from each of the three rear-end crash risk models show that the rear-end crash risk in any type of work zone increases with the heavy vehicle percentage and the lane traffic flow rate. Interestingly, it is found that the lane closer to the work zone is strongly associated with a higher risk of a rear-end crash. In addition, a truck has a much higher risk of being involved in a rear-end crash than a car. One possible reason for this is that a truck has less braking capability. Furthermore, the expressway work zone is found to have a significantly higher risk of rear-end crash occurring than the arterial work zone, consistent with the findings from previous studies.

The literature review showed that previous work zone risk assessment studies focused on separately evaluating the occurrence frequency and the severity of work zone crashes. However, this separate evaluation cannot completely reflect a facility’s risk as rendered by a broad range of accidents from frequent-minor to rare-major. In reality, traffic safety engineers are more concerned with the vehicle occupant’s casualty risk, i.e., the likelihood of a vehicle occupant being killed or injured in a work zone crash and the relationship between the frequencies and consequences of work zone crashes.

To address these concerns about the assessment of a vehicle occupant’s casualty risk, this thesis finally builds a probabilistic QRA model to evaluate the casualty risk by combining the frequencies and consequences of all accident scenarios triggered by work zone crashes. The proposed probabilistic QRA model consists of a
crash frequency estimation function, an event tree and a consequence estimation model. The event tree consists of seven intermediate events including age, crash unit, vehicle type, alcohol, light condition, crash type and severity. To allow for the parameter uncertainty, the probabilities of the intermediate events vehicle type, crash type and severity are considered to be random variables, represented by means of probability distributions.

The probabilistic QRA model successfully reports individual and societal risk in work zone. The results show that the occurrence probability of an intermediate event severity that has large statistical variation causes a big uncertainty on the individual fatality risk. This is mainly because the occurrence frequency of an occupant being killed is greatly affected by the probability of the severity event. Both the mean speed and the emergency response time (ERT) are found to have negative influences on the individual fatality risk as well as on the injury risk. More specifically individual fatality risk is found to significantly decrease with the ERT, whereas individual injury risk is insignificantly affected. Compared with the ERT, the mean speed has larger effects on both the individual fatality and injury risks, suggesting that reducing speed is a more effective method to mitigate individual fatality and injury risks.

### 9.2 Recommendations for Future Work

This research has dealt with three work zone operational issues — the
estimation of work zone traffic delay, the analysis of subwork zone operational strategy and the assessment of work zone risk. This section gives details how the work can be further improved in the future.

For the work zone traffic delay estimation problem, one limitation of the developed HCA model is that it does not take road networks into account, and its application is therefore limited. To address this problem, future studies should extend the model by incorporating a road network. In addition, the model should take into account the possible lane changing behavior that vehicles in the through lane may change into other through lanes in order to avoid merging conflicts from the closed lane.

Although the decision tree-based model provides an accurate estimate of work zone capacity, one limitation is that the tree structure and prediction accuracy might alter significantly if different strategies are applied for creating the training and checking data. To build a more stable model, future studies might focus on the estimation of work zone capacity using the ensemble tree approach, which combines all individual decision trees, based on different training and checking data, into one model. Local data should be also collected for the modeling effort.

For the subwork zone operational strategy, it has been shown that the working rate and fixed setup time of a subwork zone can significantly affect the optimal solution. In reality, these two factors are always changing according to the traffic conditions. Hence, it would be interesting to formulate them as stochastic variables for the determination of subwork zone operational strategy. Since this research only
considers the situation where there is only one work zone project in a given road segment, future research should also attempt to optimize the subwork zone operational strategies for multiple work zone projects in a road network.

The crash risk assessment study is limited to assessing work zone rear-end crash risk at the work zone activity area because of data limitations. Without taking into account lane changes, this study is not applicable for estimating the work zone rear-end crash risk in the work zone advance warning and transition areas, since a large number of lane change maneuvers occur in these areas. Hence, future research is needed to assess the rear-end crash risk by taking into lane changes at work zone advance and transition areas. Whether lane changing behavior could reduce a vehicle’s rear-end crash risk would also be an interesting future research topic.

For the work zone casualty risk assessment, this research only considers the casualty risk caused by work zone crashes, meaning that the risk is underestimated since some other top events such as fires, which can also result in casualties at work zones, are ignored. To eliminate this problem and improve the accuracy of the assessment, a future study should take into account all possible top events in the QRA model. In addition, risk control measures should be incorporated, in order to determine the most economic and effective measures of the mitigating of work zone casualty risk.
REFERENCES


Al-Kaisy, A. F., Hall, F. L., 2001. Examination of effect of driver population at

212
freeway reconstruction zones. *Transportation Research Record 1776*, TRB, pp.35-42.


Bham, G.H., Benekohal, R.F., 2004. A high fidelity traffic simulation model based on


Practice, 28(5), pp.433-444.


Garber, N.J., Zhao, M., 2002. Distribution and characteristics of crashes at different work zone locations in Virginia. Transportation Research Record 1794, TRB, pp.19-25


Kallberg, V.P., 1998. The two effects of speed on accidents: number and severity. 


NUREG, 1975. WASH 1400 (NUREG-75/014), Reactor Safety Study, Washington, D.C.


APPENDIX: CURRICULUM VITAE

Education

2002-2006 Bachelor, School of Engineering, Sun Yat-Sen University, China
2007-2011 PhD, Department of Civil and Environmental Engineering, National University of Singapore, Singapore

Awards

1. The EASTS (Eastern Asia Society for Transportation Studies) Best Paper Award For Methodological Development (2011)

2. The TRB (Transportation Research Board) AHB40 Best Paper Nomination Award (2011)

List of Journal Papers


