MD-06-SP508B4C

Robert L. Ehrlich, Jr., *Governor*
Michael S. Steele, *Lt. Governor*

Robert L. Flanagan, *Secretary*
Neil J. Pedersen, *Administrator*

STATE HIGHWAY ADMINISTRATION

RESEARCH REPORT

Optimization of Work Zone Decisions through Simulation

PAUL SCHONFELD
NING YANG, SHIN-LAI TIEN
UNIVERSITY OF MARYLAND

FINAL REPORT

July 2006

Technical Report Documentation Page

1. Report No. **MD-06-SP508B4C**	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle **Optimization of Work Zone Decisions Through Simulation**		5. Report Date **July 2006** 6. Performing Organization Code
7. Author/s **Paul Schonfeld, Ning Yang, Shin-Lai Tien**		8. Performing Organization Report No.
9. Performing Organization Name and Address **University of Maryland** **Department of Civil and Environmental Engineering** **College Park MD 20742**		10. Work Unit No. (TRAIS) 11. Contract or Grant No. **SP508B4C**
12. Sponsoring Organization Name and Address **Maryland State Highway Administration** **Office of Policy & Research** **707 North Calvert Street** **Baltimore MD 21202**		13. Type of Report and Period Covered **Final Report** 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Highway maintenance, especially pavement rehabilitation or resurfacing, requires lane closures, which can greatly affect traffic performance and traffic safety due to the reduction in road capacity. The objective of this research is to develop a comprehensive work zone evaluation and decision support tool for highway maintenance planning and traffic management plan development.
In the first part of this report, a work zone cost model is developed based on three approaches: (1) an analytic method for steady traffic inflows; (2) an analytic method for time-dependent traffic inflows; and (3) a simulation method, which uses CORSIM, a widely-used simulation program, to evaluate work zone conditions in a user-defined roadway network.
In the second part of this study, work zone optimization models are developed based on the above three methods. When using the analytic method for steady traffic inflows, a closed-form minimized total cost can be obtained through differential calculus. When the analytic method for time-dependent traffic inflows or the simulation method is used, a newly-developed heuristic algorithm, named two-stage modified simulated annealing (2SA), is employed for optimization.

17. Key Words **Work Zone, Cost, Optimization**	18. Distribution Statement: No restrictions **This document is available from the Research Division upon request**.		
19. Security Classification (of this report) **None**	20. Security Classification (of this page) **None**	21. No. Of Pages	22. Price

Form DOT F 1700.7 (8-72) Reproduction of form and completed page is authorized.

EXECUTIVE SUMMARY

This report is the product of a project entitled "Optimization of Work Zone Decision through Simulation" conducted at the University of Maryland, College Park, under the sponsorship of the Maryland State Highway Administration (SHA). The objective of this research is to develop a comprehensive work zone evaluation and decision support tool for supporting highway maintenance planning and traffic management plan development. This tool may be used to evaluate different work zone plans as well as to optimize various decisions on work zones characteristics and traffic control plans in order to minimize the combined total costs for highway agencies and users. In this study, two analytic methods and a simulation method are developed to calculate the work zone costs with different work zone characteristics for the work zone plan evaluation and optimization. Users may choose to use the analytic methods or the simulation method, depending on the availability of data, the level of detail desired for the analysis and the allowable running time. A software package incorporating proposed analysis methods is developed and a users' guide for the software package is provided in the appendix to this report.

Highway maintenance, especially pavement rehabilitation or resurfacing, requires lane closures, which can greatly affect traffic performance and traffic safety due to the reduction in vehicle capacity. Good decisions on work zone characteristics, such as lane closure strategy, scheduling, work zone configuration, work rate and traffic control strategy, can significantly increase the work efficiency and safety as well as decrease the negative impacts of traffic disruption.

In the first part of this work, a work zone cost model is developed based on three approaches: (1) an analytic method for steady traffic inflows; (2) an analytic method for time-dependent traffic inflows; and (3) a simulation method, which uses CORSIM (which is short for "Corridor Simulation"), a widely-used simulation program, to evaluate work zone conditions in a user-defined roadway network. From case studies, we find that CORSIM estimates higher delays than the analytic methods under uncongested traffic conditions and lower delays than the analytic methods under congested conditions. This can be explained by the inability of CORSIM to calculate the delays of the vehicles that cannot enter the network as the queues spill back beyond traffic entry nodes in an over-saturated road network.

In the second part of this study, work zone optimization models are developed based on the above three methods. When using the analytic method for steady traffic inflows, a closed-form formulation of the total cost can be obtained. Classic optimization methods using differential calculus are applied to identify the preferred solutions.

When the analytic method for time-dependent traffic inflows or the simulation method is applied, we have no simple expressions for the objective function in terms of the decision variables. Therefore, a heuristic optimization algorithm, named two-stage modified simulated annealing (2SA), is developed to search in solution space for an optimized solution. Optimization models based on the analytic method for time-dependent traffic inflows (A2SA) or the simulation method (S2SA) are proposed and tested through numerical examples. In order to reduce the computation time and effort while keeping a desirable precision level, a hybrid approach (H2SA), is proposed. In the two stages of the optimization algorithm, the analytic method is applied in the *Initial Optimization* step and the simulation method is used in the second *Refined Optimization* step. The numerical experiment shows that H2SA can obtain optimized results close to S2SA.

ACKNOWLEDGEMENTS

The researchers wish to thank the Office of Traffic and Safety and Office of Policy and Research of the Maryland State Highway Administration for supporting this project. They also wish to individually thank Ms. Allison Hardt and Mr. Jawad Paracha for their great help with extensive technical support, advice and feedback.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Introduction

1.1 Background

The prosperity and economic growth experienced in the United States during the 1990s contributed to an increase in demand for many modes of surface travel. However, a large fraction of the nation's current transportation infrastructure has reached the end of its design life. The deterioration of the national highway system severely affects wear-and-tear on vehicles, fuel consumption, travel time, congestion, and comfort, as well as public safety. To maintain the highway infrastructure while traffic continuously increases, state and federal transportation agencies have shifted their focus from building new highways to maximizing the performance of existing surface transportation systems. Highway maintenance and reconstruction activities are likely to increase in number, duration, and scope in the near future.

In order to perform such activities on roadways, segments of lanes and shoulders are sometimes closed to form work zones. This can greatly affect traffic performance and traffic safety since vehicle capacity is reduced.

Good decisions about work zone characteristics, such as lane closure scheduling, work zone configuration, work rate and traffic control strategy, can significantly increase the work efficiency and safety as well as decrease the negative impact of traffic disruption. The total cost, including both agency and user cost, is a very useful and appropriate measure for evaluating the work zone decisions. It is worthwhile to develop appropriate work zone analysis methods which can aid highway agencies in developing cost-effective highway maintenance or reconstruction plans.

1.2 Research Objective

The objective of this research is to develop a comprehensive work zone evaluation and decision support tool for supporting highway maintenance planning and traffic management. This tool should be usable not only for evaluating different work zone plans but also for automatically searching for optimal solutions to the following questions:

(1) What are the best construction periods, including starting times, pauses and ending times, given time constraints?

(2) What are the most advantaged work zone lengths?

(3) What kind of work zone configuration is preferable, (e.g. number of closed lanes, crossover), depending on circumstances?

(4) How much traffic (if any) should be diverted to other routes if such detour routes are available?

(5) What should be the proper work rate and the corresponding work cost?

A user-friendly software package is also developed to implement the methods developed in this research project.

1.3 Research Scope and Tasks

The research issues covered in this study can be considered on two levels. The first level is the evaluation of the work zone impacts given a certain set of input variables. Based on the performance measurement model formed in the first level, the optimization of the input parameters can be performed in the second level. Corresponding to these two levels, two analysis tools are provided in this project: (1) a Scenario Comparison Tool and (2) a Decision Optimization Tool.

Based on the methods of the analysis, the scope of this study covers (1) analytic models and (2) a simulation model for evaluating work zone impacts. Based on highway configuration, the scope of this study covers (1) two-lane two-way highway work zones and (2) multiple-lane two-way highway work zones. Based on traffic flow patterns, the scope covers (1) steady traffic inflows and (2) time-dependent inflows. Based on detour type, the methods cover three cases: (1) no detour, (2) a single detour and (3) multiple detour paths. The correlations between different scope categories are displayed in Figures 1.1 and 1.2.

The following tasks have been completed in this research project:

- Identification of characteristics of work zone operations and specification of important work zone decisions.
- Determination of performance measurements to evaluate work zone impacts.
- Development of performance measurement model for given work zone characteristics based on analytic method.
- Development of performance measurement model for given work zone characteristics based on simulation method.

- Development of optimization model for work zone decisions. The subtasks include: The specification of the scope of the optimized work zone decisions, the formulation of the objective function and the development of the optimization algorithm.

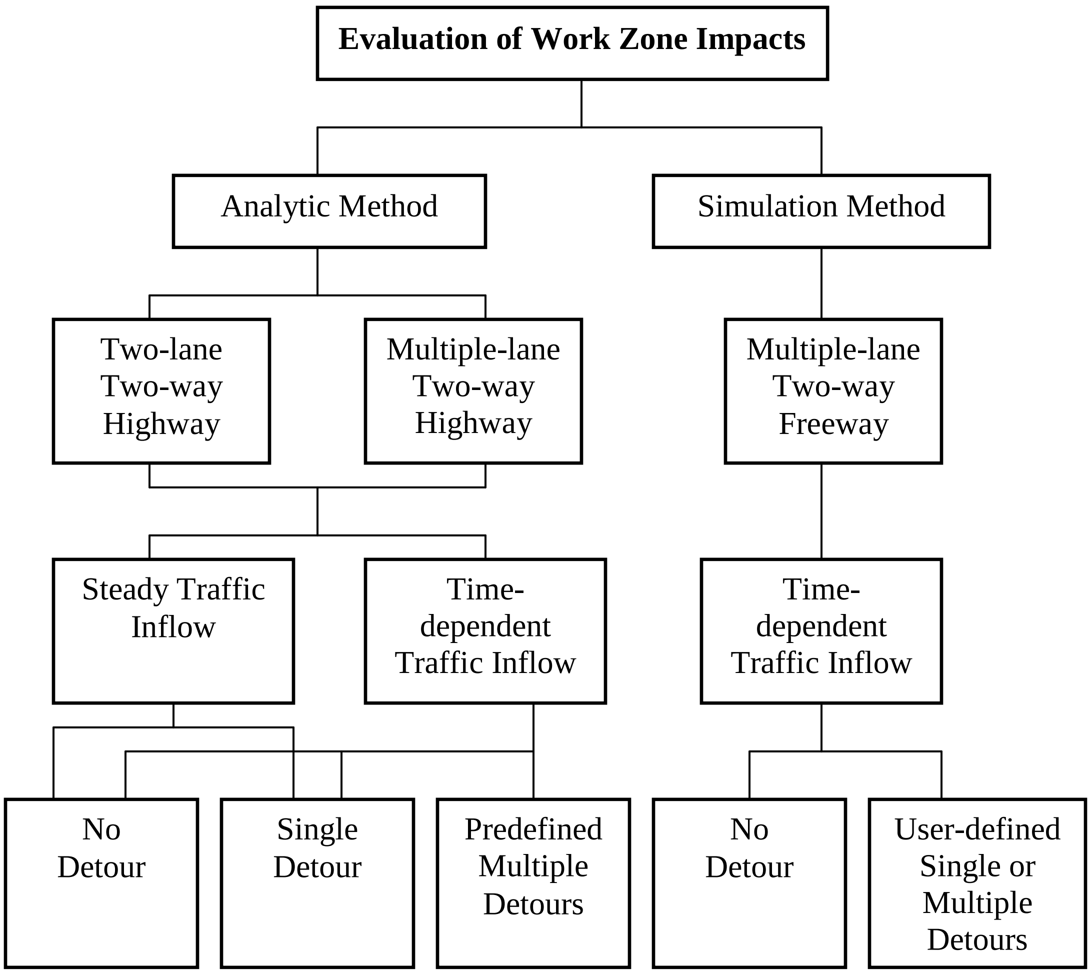


Figure 1.1 Scope of Level 1

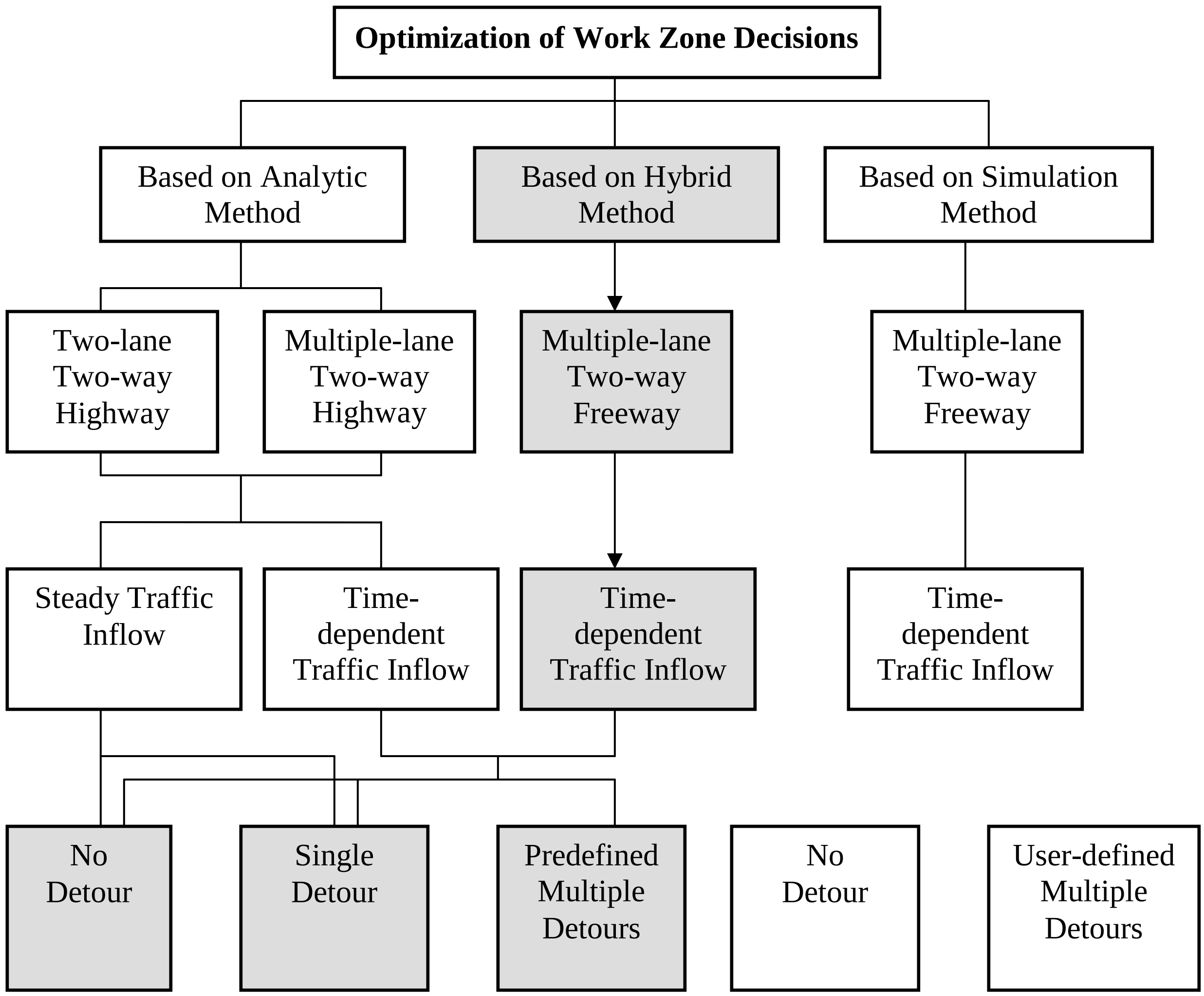


Figure 1.2 Scope of Level 2

1.4 Technical Approaches

The total cost (C_T) spent on completing per lane mile maintenance work is used as the major performance measurement to evaluate a work zone plan consisting of certain work zone characteristics, such as work zone lengths, work zone configuration. The total cost can be expressed as a function of work zone characteristics:

$$\begin{aligned} C_T &= C_T \text{ (work zone characteristics)} \\ &= C_M \text{ (work zone characteristics)} + C_U \text{ (work zone characteristics)} \end{aligned} \tag{1-1}$$

where C_T is total cost, C_M is maintenance cost, or supplier cost, and C_U is user cost. The controllable variables affecting C_M include work zone length, fixed setup cost, and average maintenance cost per unit length; the controllable variables affecting C_U include work zone length, traffic volumes, speed, diverted fractions (if detours are available), etc.

The user cost mainly depends on the users' time value and the delay times caused by work zones. In this study, two approaches are applied to estimate the user delays. One is an analytic approach, which includes an analytic method for steady traffic inflows and an analytic method for time-dependent traffic inflows. The other is a simulation approach, which uses CORSIM, a commercial simulation program, to simulate work zone conditions.

The objective of the work zone optimization problem is to find the optimal work zone plan which can minimize the total cost for the maintenance work. The objective function for work zone activities can be expressed as follows:

$$\text{Min } C_T = C_M + C_U \tag{1-2}$$

The decision variables are major controllable variables affecting the costs. Some constraints such as work time constraints are considered.

When the user cost is estimated with the analytic method for steady traffic inflows, a closed-form equation for the total cost can be obtained. Classic optimization methods based on differential can be used to find the optimum solutions.

When the user cost is obtained from the analytic method for time-dependent traffic inflows or the simulation method, we have no simple expressions for the objective function in terms of the decision variables. Therefore, a heuristic optimization algorithm, named two-stage modified simulated annealing algorithm (2SA), is developed to search the solution space and find an optimized solution.

1.5 Organization of Report

To achieve the above research purposes, this report consists of the following nine chapters. The interrelations among these chapters and their development sequence are shown in Figure 1.3.

The focus of each chapter is detailed below.

Chapter 1, "Introduction," contains background information, research objectives, and the technical approach used in this study.

Chapter 2, "Literature Review," focuses on reviewing research performed in the previous ten years that is considered relevant to the project objectives.

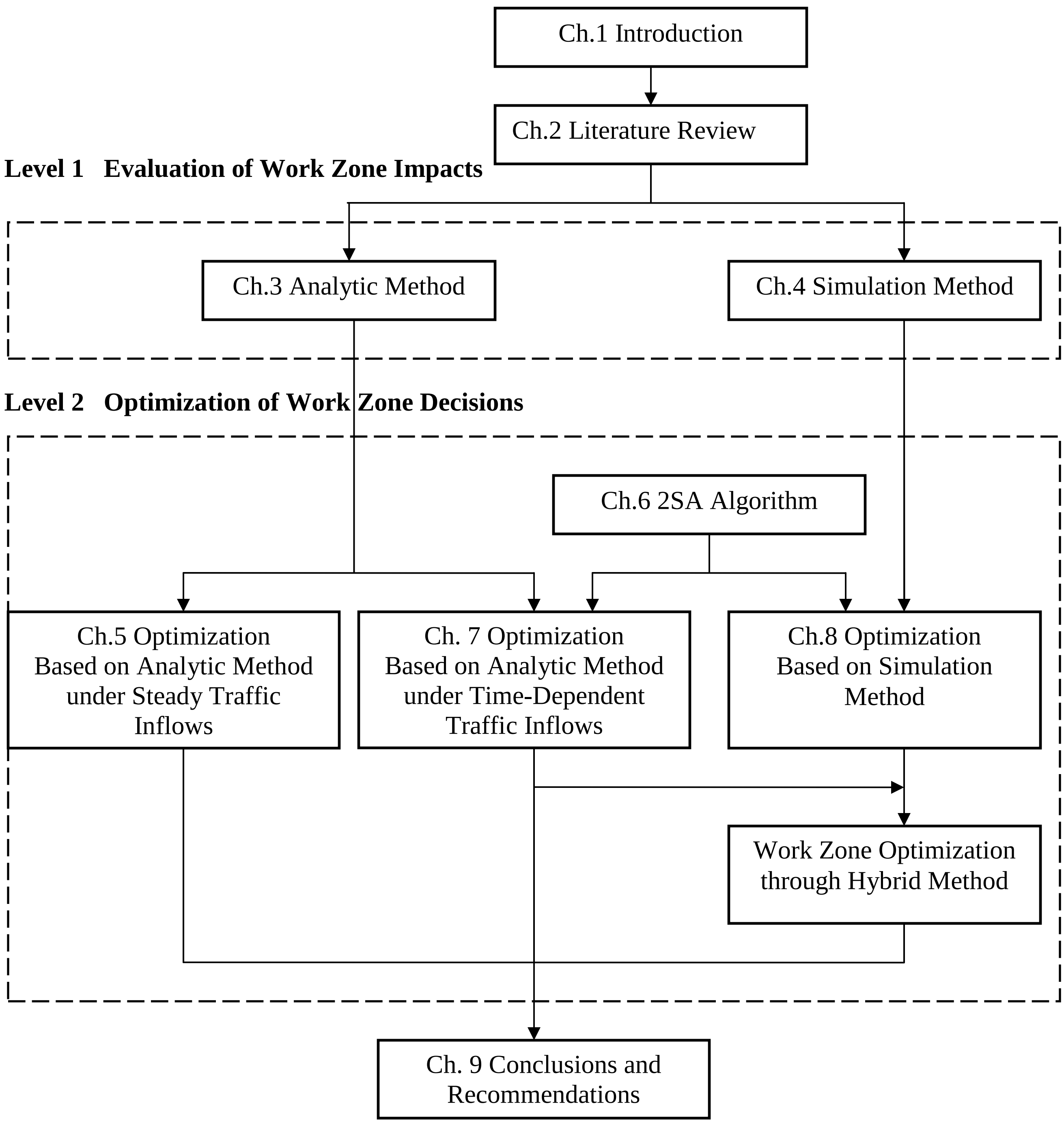


Figure 1.3 Interrelations among Chapters

In Chapter 3, "Analytic Model of Work Zone," analytic models are developed to evaluate the traffic impact caused by work zone lane closures. Both steady traffic inflows cases and time dependent traffic inflows cases are discussed.

In Chapter 4, “Simulation Model of Work Zone,” simulation method is applied to evaluate work zone impacts. Work zone conditions are simulated using a microscopic simulation program, CORSIM.

Chapter 5, “Work Zone Optimization based on Analytic Model for Steady Traffic Inflows,” presents an optimization method for work zones under steady traffic volumes. Tradeoffs between construction time and cost are considered.

The purpose of Chapter 6, “Two-Stage Modified Simulated Annealing Algorithm,” is to introduce an optimization algorithm developed for work zone optimization based on the analytic method for time-dependent traffic inflows or the simulation method.

Chapter 7, “Work Zone Optimization based on Analytic Model for Time-dependent Traffic Inflows,” applied the optimization algorithm developed in Chapter 6 to the optimization based on the analytic method for time-dependent traffic inflows.

Chapter 8, “Work Zone Optimization based on Simulation Model,” applied the optimization algorithm based on the simulation method. A hybrid method, labeled H2SA, in which both the analytic model and the simulation model are used to evaluate work zone solutions in the optimization process, is also developed. A case study is used to compare the performances of optimization based on analytic methods, optimization based on simulation methods and optimization through hybrid method.

Chapter 9, “Conclusions and Recommendations,” summarizes the significant research results of this study. Recommendations for further research are also discussed.

1.6 Summary

In this chapter, after the introduction of the background of this study, the research objectives are presented. Then we discuss the scopes covered by this project and the research tasks. Approaches to accomplish the tasks are briefly introduced.

Chapter 2 Literature Review

The purpose of this chapter is to concisely review the relevant literature on the operation, management and optimization of highway work zones. This chapter is organized into 4 major parts: (1) work zone operation and management, (2) work zone cost estimation (3) delay estimation at work zones, and (4) work zone optimization.

To evaluate the impact caused by work zones, the basic characteristics of work zones should be defined. Those characteristics are discussed under work zone operation and management issues.

In quantifying the impacts of a work zone, the most commonly considered factors are: (1) traffic delay and safety, (2) project cost, (3) constructability and (4) environmental impact (Martinelli and Xu, 1996). User costs are included in traffic delay and safety. Among the above factors, the first two factors are the most widely used.

It is natural for researchers to relate the design of work zone characteristics with the work zone impact evaluation through optimization methods. Those studies are briefly reviewed in this chapter.

2.1 Work Zone Operation and Management

A work zone is defined in the Highway Capacity Manual as an area of a highway in which maintenance and construction operations are taking place that impinges on the number of lanes available to traffic and affect the operational characteristics of traffic flowing through the area. To evaluate the impact caused by work zones, the characteristics of the work zones must be specified. Work zone characteristics of concern include such factors as work zone length, number and capacity of lanes open, duration of lane closures, timing of lane closures, posted speed, and the availability and traffic characteristics of alternative routes.

Work zone length is an important issue that has been relatively neglected. In general, longer zones tend to increase the user delays, but the maintenance activities can be performed more efficiently (i.e., with fewer repeated setups) in longer zones (Schonfeld and Chien, 1999). In practice, such lengths have been usually designed to reduce costs to highway agencies rather than users.

The lane closure type is one of the major factors which affect the vehicle capacity in work zones and it also affects agency costs to a considerable degree. There are three main lane closure types for work zones: partial lane closure, full lane closure and crossover (Pal and Sinha, 1996). In a partial lane closure one or more lanes are closed in one or both directions, but not all the lanes in one direction are closed simultaneously. Traffic cones, drums, or concrete barriers are used to close the lanes, and maintenance and rehabilitation activities are performed on the closed lanes. During full road closure, traffic is detoured, allowing full access to roadway facilities. Under the appropriate conditions, a full closure can be an effective way to complete projects with shorter duration and less safety risks. Departments of transportation in Oregon, Kentucky, Michigan, Ohio, Washington, and Delaware have experience in using a full closure approach to conduct road rehabilitation/reconstruction projects (FHWA-OP-04-009, 2003). In a crossover arrangement, all lanes in one direction are closed and both direction of traffic is brought to one side of a highway. Due to the additional cost of constructing the crossover facility, the fixed set up cost in cases of crossover is always higher than in cases of partial lane closure at sites. However, in a crossover lane closure strategy sufficient working spaces are available, which may improve the safety of the workers and increase their productivity as well as the quality of their work. It is noted that sometimes closed lanes may include not only maintained lanes, but also additional lanes which are used to provide access to and from the work site for maintenance and construction vehicles or provide buffer space to separate traffic and work sites from safety consideration.

Since travel demands are time-varying, work zone scheduling can greatly affect the traffic impact caused by lane closures. Work zones can be categorized into three designations: (1) short-term sites, at which maintenance work lasts less than one day (24 hours) (Jiang, 2003); (2) intermediate sites, at which work lasts over one day but less than four days; (3) long-term sites, at which work lasts more than four days (Rouphail, 1988). Unlike in long-term projects which continuously occupy the road space for several days or months, short term and intermediate work zones are often limited to the time defined in some construction windows, e.g. off-peak daytime, nighttime periods, or weekend periods, in order to avoid the higher volume daytime hours and associated traffic delays.

2.2 Work Zones Cost Estimation

Work zone costs may be classified into two categories: (1) agency costs and (2) user costs.

Agency costs are those expenses required to finish the work zone activities based on the work types. Those normally include labor costs, equipment costs, material costs and traffic maintenance costs. Underwood (1994) analyzed the work duration and the maintenance cost per 10,000 m^2 for five different roadway maintenance activities (i.e., surface dressing, asphalt surface, porous asphalt, 10% patching, and milling out). The average maintenance costs were calculated based on prices quoted to highway authorities in the summer of 1993.

User costs also receive great attention in work zone analysis because they tend to dominate other costs and because community concerns and reactions to work zone activities affect many aspects of work zone decisions. User costs are usually evaluated considering at least three components: (1) travel time delay costs, (2) additional vehicle operating costs and (3) work zone related accident (crash) costs (Najafi and Soares, 2001; USDOT/FHWA, 1996).

Delay costs result from increases in travel time through the work zone-from speed reduction, congestion delays, or increased distances as a result of taking a detour (USDOT/FHWA 1989). Typically, the delay cost can be determined by multiplying the user delay by the value of user time (Wall, 1998). Studies on user delay estimation will be introduced in the next subsection.

Vehicle operating costs are the costs associated with owning, operating, and maintaining a vehicle including: fuel consumption, tire wear, maintenance and repair, and so on. Many factors such as vehicle characteristics, vehicle speed, road geometrics, road surface type and condition, environmental factors can affect vehicle operating costs. Vehicle operating costs can be formulated empirically or mechanistically, deterministically or probabilistically. In many studies, vehicle operating costs were formulated using classical regression analysis of historic information (Booz Allen & Hamilton, 1999; Berthelot and etc., 1996; etc.).

Accident (crash) costs are related to the historical crash rate, vehicle miles of travel, delay, work zone configuration, and average cost per crash. Crash rates are commonly specified as crashes per 100 million vehicle miles of travel (100 M VMT). Overall crash rates for the various functional classes of roadway are fairly well established. Crash rates for work zones, however, are not easy to estimate due to the limited amount of data and the variety of work zone types. McCoy et al. (1980) found the average rate was 30.8 crashes per 100 million vehicle miles (acc/100 mvm) on I-80 in Nebraska between 1978 and 1984. Pigman and Agent (1990) found that the work zone crash rate varied from 36 to 1,603acc/100 mvm on different highways based on the crash data collected from the Kentucky Accident Reporting System (KARS) for the 1983-1986 periods. Chien and Schonfeld (2001) determined work zone crash cost from the crash rates multiplied by the product of the user delay and average cost per crash and then divided by work zone length.

2.3 Work Zone Delay Estimation

The delays related to work zones can be classified into five categories (Jiang, 1999; NJDOT Road User Cost Manual, 2001): (1) deceleration delay by vehicle deceleration before entering a work zone, (2) moving delay by vehicles passing through work zones with lower speed, (3) acceleration delay by vehicle acceleration after exiting work zone, (4) queuing delay caused by the ratio of vehicle arrival and discharge rates, and (5) Detour Delay by the additional time necessary to traveling the excess distance the detour imposes.

Over the years a number of manual and computerized approaches have been developed for estimating the work zone delays (McCoy and Peterson, 1987; Schonfeld and Chien, 1999; Venugopal and Tarko, 2000; Chien and Schonfeld, 2001; and Chien et al., 2002; etc.).

2.3.1 Analytic Method

2.3.1.1 Delay Models

Two well-known methods are widely used to analyze queuing delays caused by bottleneck: (1) the deterministic queuing models (Abraham and Wang, 1981; Dudek and Rechard, 1982; Morales, 1986; Schonfeld and Chien, 1999) and (2) the shock wave models (Richard, 1956; Wirasinghe, 1978; Al-Deek et al., 1995; etc.).

The deterministic queuing analysis is recommended by the Highway Capacity Manual (HCM) as the standard delay estimation technique for freeway zones (TRB 1994). It is essentially a graphical procedure using a deterministic queuing diagram with the x-coordinate as time and the y-coordinate as the cumulative number of vehicles.

In the shockwave model, the traffic flow is assumed to behave like a fluid, and a backward shock wave develops when demand exceeds capacity. This model is often used to estimate incident congestion. However, the shock wave speed is approximated based on traffic density, which is often difficult to measure or estimate.

2.3.1.2 Computerized Software

QUEWZ and QUICKZONE are the most used software packages for estimation of queue lengths and delays in work zones. Both of these software packages model traffic flow at a macroscopic level.

The computer model, called Queue and User Cost Evaluation of Work Zone (QUEWZ), was developed by Memmott and Dudek (1984) to assess work zone user costs. The most recent upgrade version is QUEWZ-98. It analyzes traffic conditions on a freeway segment with and without a lane closure in place and provides estimates of the additional road user costs and of the queuing resulting from a work zone lane closure. The road user costs calculated include travel time, vehicle operating costs, and excess emissions. That model does not consider any alternate path and the effect of diverting traffic to it.

QuickZone is a work zone delay estimation program developed in Microsoft Excel. The primary functions of QuickZone include quantification of corridor delay resulting from capacity decreases in work zones, identification of delay impacts of alternative project phasing plans, supporting tradeoff analyses between construction costs and delay costs, examination of impacts of construction staging, by location along mainline, time of day (peak vs. off-peak) or season, and assessment of travel demand measures and other delay mitigation strategies. The costs can be estimated for both an average day of work and for the whole life cycle of construction. The Maryland State Highway Administration and the University of Maryland (Kim and Lovell, 2001) used QuickZone's open source code to customize the program to meet the State's needs. The University has added its own capacity estimation model to the program and has used a 24-hour traffic count, instead of the average daily traffic count found in original version. However, this program requires the users to input a great deal of information concerning a particular project, which may discourage the application of the software. To use the QuickZone program the user must first create a network of traffic facilities and then input hourly traffic volumes and capacities of all the links.

2.3.2 Simulation

Although the concept of deterministic queuing model is widely accepted by practitioners for estimating queuing delay, the delay is usually underestimated (Mashine and Ross, 1992; Nam and Drew, 1998; Chien and Chowdhury, 2000; Najafi and Soares, 2001) due to the neglected approaching and shock-wave delays. Besides, for a complex road network, analytical methods may not estimate user delays precisely.

As valuable analysis tools, microscopic traffic simulation models have been applied in various problems in work zone studies, such as the evaluation of traffic management plans, estimation of capacity and queue length, and optimization of traffic controls (Nemeth and Rathi, 1985; Cohen and Clark, 1996; Chien and Chowdhury, 1998; Maze and Kamyab, 1999; Schrock and Maze, 2000; Lee, Kim and Harvey, 2005; etc.). CORSIM (including NETSIM and FRESIM), VISSIM, PARAMICS and INTERGRATION are among the most widely used microscopic simulation models.

Simulation models can output different measures of effectiveness (MOE's). In work zone analysis, delay, travel time, speed and volume are frequently used MOEs. However, simulation can be quite costly in terms of both computer and analysis time. Advanced computer and parallel processing techniques can be useful to decrease the simulation time. The combination of analytic method and simulation method is also explored. Chien and Chowdhury (2000) developed a method to approximate delays by integrating limited amounts of simulation data and the concept of deterministic queuing model. In their study, simulation is applied to estimate average queuing delay with various ratios of entry volume to work zone capacity.

Only a few studies have been performed to date to validate the use of simulation models for work zone applications.

Dixon et al. (1995) evaluated the suitability of FRESIM for a simple freeway lane closure by comparing simulated behavior to the observed behavior of a study site. They concluded that FRESIM provided similar results to those observed in the field for low volume conditions. However, high volume conditions were not accurately simulated.

Middleton and Cooner (1999) evaluated three simulation models, CORSIM, FREQ and INTEGRATION, for simulating congested freeways. The calibration and validation performances of those models were tested using data collected on Dallas freeways. They concluded that all of the three models performed relatively well for uncongested conditions;

however, the performance became sporadic and mostly unreliable for congested conditions. The CORSIM model had the best overall performance, compared with the other two models.

Chitturi and Benekohal (2003) compared the queue length measured from field data to the results from FRESIM, QUEWZ, and QuickZone Software. They found that the results generated by QUEWZ did not match the field data. FRESIM either underestimated or overestimated the queue lengths. QuickZone underestimated the queue lengths generally.

2.4 Work Zone Optimization

A lot of efforts have been devoted to optimizing work zone decisions to minimize negative impacts, usually measured by the total cost. McCoy et al. (1980) provided a simple framework to find the optimum work zone length by minimizing the total cost including construction, user delay, vehicle operating, and crash cost in construction and maintenance zones of rural four-lane divided highways. The user delay costs were modeled based on average daily traffic (ADT) volumes, while the crash costs were computed by assuming that the crash rate per vehicle mile was constant in a work zone area. The optimal work zone length was derived based on 1979 data. Because the unit cost factors had changed considerably since 1981, McCoy and Peterson (1987) found the optimum work zone lengths to be about 64% longer that those used previously.

Martinelli and Xu (1996) added the vehicle queue delay costs into McCoy's (1980) model. The work zone length was optimized by minimizing the total user cost, excluding the maintenance and crash costs. Viera-Colon (1999) developed a similar model of four-lane highways which considered the effect of different traffic conditions and an alternate path. However, that study did not develop alternative selection guidelines for different traffic flows or road characteristics. Schonfeld and Chien (1999) developed a mathematical model to optimize the work zone lengths plus associated traffic control for two-lane, two-way highways where one lane at a time is closed under steady traffic inflows. They found the optimal work zone length and cycle time for traffic control and minimized the total cost, including agency cost and user delay cost. No alternative routes were considered in that study. They (2001) then developed another model to optimize the work zone length on four-lane highways using a single-lane closure strategy. Based on the previous work, Chen and Schonfeld (2005a, 2005b) developed work zone length optimization models for two-lane and four-lane highway with a single alternate route under steady traffic inflows.

Fwa, Cheu, and Muntasir (1998) developed a traffic delay model and used genetic algorithms to optimize scheduling of maintenance work for minimizing traffic delays subject to constraints of maintenance operational requirements. Pavement sections, work teams, start time and end time for each section were scheduled. Other conditions in that study were given, e.g. work zone configuration and available work duration for each team, and road section length. These variables were not optimized in that study. Chang, Sawaya, and Ziliaskopoulos (2001) used traffic assignment approaches to evaluate the traffic delay, which include the impact of work zone combinations on an urban street network. A tabu search methodology was employed to select the schedule with the least network traffic delays.

Chien, Tang, and Schonfeld (2002) developed a model to optimize the scheduling of work zone activities associated with traffic control for two-lane two-way highways where one lane at a time is closed considering time-varying traffic volumes during four periods in a day: morning peak, daytime, evening peak, and nighttime periods. A greedy method is used as the search approach. Jiang and Adeli (2003) used neural networks and simulated annealing to optimize work zone lengths and starting times for short-term freeway work zones using average hourly traffic data, considering factors such as darkness and numbers of lanes. More complete scheduling plans for multiple-zone maintenance projects were not attempted in that work. Chen and Schonfeld (2004) developed a set of models to simultaneously optimize the work zone length, scheduling, lane closure strategy and diversion fractions, using simulated annealing search algorithm. Two-lane and multiple-lane highways, single and multiple detours as well as steady and time-varying traffic volume, are considered in their models.

All the above studies used macroscopic analytical methods (e.g. deterministic queuing analysis) to estimate user delays. However, analytical methods are based on some simplified assumptions which lead to the neglect of detailed representations. Therefore analytical methods may not be able to provide satisfactory solutions for complex transportation networks.

With the increasing development of computing technology, simulation-based optimization has received considerable attention. This process seeks to find the best value of some decision variables for a system where the performance is evaluated based on the output of a simulation model of this system (Olafsson and Kim, 2002).

Cheu et al. (2004) presented a hybrid methodology to schedule maintenance activities at various sites in a road network, using genetic algorithm (GA) as an optimization technique, coupled with a traffic-simulation model to estimate the total travel time of users. This study demonstrated the availability of simulation-based optimization technology in work zone problems, although it did not focus on work zone length and duration optimization problem.

2.5 Summary

After a review of the above studies, it appears that many methods, consisting of both analytic methods and simulation models, have been developed and applied in various problems in work zone studies. Some analytical and heuristic methods were proposed for solving work zone optimization problems. However, most of the previous studies are based on analytic methods, which may not be precise due to over-simplified assumptions especially in complex traffic network. Few of studies integrate simulation with optimization. A main barrier is that simulation is a very time-consuming way to evaluate the objective function in an optimization process.

Chapter 3 Analytic Model of Work Zone for Steady Traffic Inflows

In this chapter, analytic models to evaluate agency cost, delay cost and accident cost are developed for two-lane highway and four-lane highway work zones under steady traffic inflows and time-dependent traffic inflows. The highway system and the characteristics of various work zone alternatives are defined in Section 3.1. In Section 3.2, the analytic model under steady traffic inflow is described. Section 3.3 presents the analytic model under time-dependent traffic inflow. Finally, numerical examples for two-lane and four-lane highways are shown in Sections 3.4.

3.1 Highway System and Work Zone Characteristics

In this study highway types are classified into two-lane two-way highways and multiple-lane two-way highways.

Two-lane two-way highways often require closing one lane for a work zone (Figure 3.1). In such circumstances, vehicles travel in the remaining lane along the work zone, alternating direction within each control cycle. Such a two-lane work zone can be treated as a one-way traffic control system in which queuing and delay processes are analogous to those at two-phase signalized intersections.

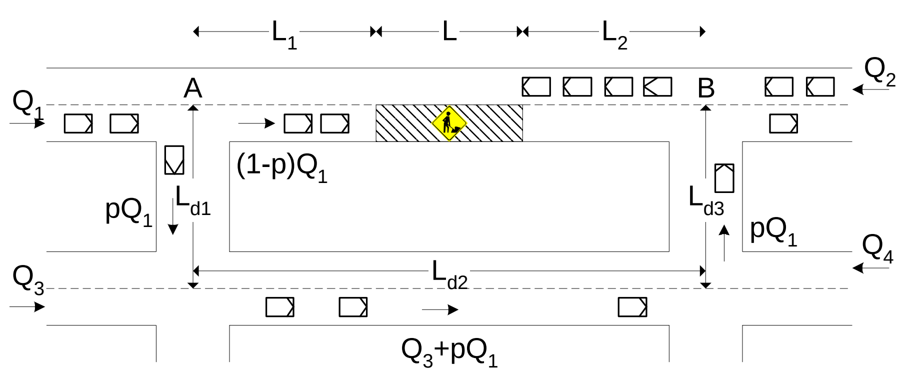


Figure 3.1 Work Zone on a Two-Lane Two-Way Highway

Pavement maintenance on multiple-lane two-way highways often requires closing one or more lanes to set up a work zone (Figure 3.2). This does not require alternating one-way control as in a two-lane highway work zone because at least one lane is usually still available for traffic in each direction. Because work zones in two-lane highways and multiple-lane highways have different delay and queuing patterns, the work zone cost functions are separately developed.

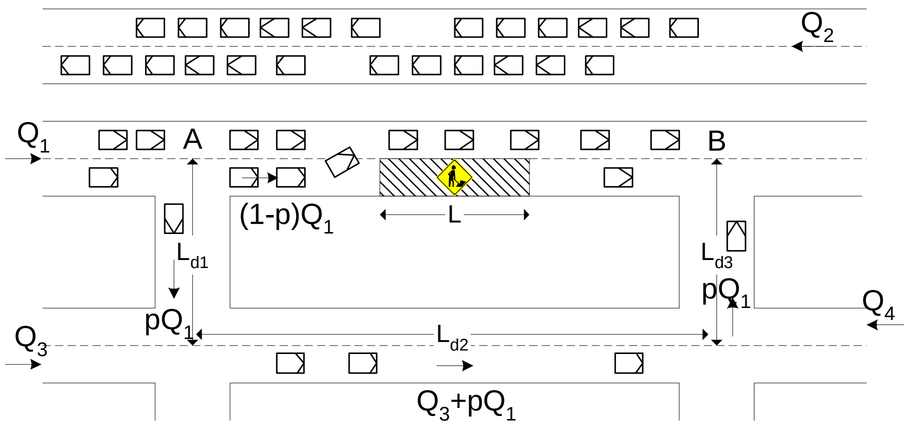


Figure 3.2 Work Zone on a Multiple-Lane Two-Way Highway

The characteristics used to describe a work zone are categorized into four categories: work zone lane closure alternatives, work zone operation characteristics, work zone rate parameters and detour types.

(1) Lane Closure Alternatives

From the review of work zone characteristics in Chapter 2, work zone lane closure alternatives are specified using the following variables given the number of total lanes in both directions (n_1 and n_2) on highway with work zones:

(1.1) The number of closed lanes in the original direction on work zone link $\boldsymbol{n_{c1}}$;

(1.2) The number of closed lanes in the opposite direction on counter work zone link $\boldsymbol{n_{c2}}$.

For two-lane two-way highways, n_{c2} is a binary variable. If the lane in the opposite direction is open for traffic in both directions, n_{c2}=0. If the opposite lane is also closed, n_{c2}=1 and in such cases the traffic flows in both directions have to be fully diverted to detour routes.

For multiple-lane two-way highways, $\boldsymbol{n_{c2}}$ represents the number of usable counter flow lanes for the traffic along the original direction when crossover strategy is applied. Therefore, $0 \leq n_{c2} \leq n_2$.

(1.3) The number of access lanes in the work zone area $\boldsymbol{n_{a1}}$;

Access lanes are lanes which are closed for providing access for demolition and construction activities. If no access lanes are needed or road shoulders are used as access lanes, $\boldsymbol{n_{a1}}$ =0.

Based on the above three variables, the number of open lanes for the traffic in both directions (n_{o1} and n_{o2}) along the work zone area and the number of maintained lanes (n_w) in a work zone can be derived from the following equations:

For two-lane two-way highways,

$$n_{o1} = n_1 - n_{c1} + (1 - n_{c2}) \qquad (3\text{-}1)$$

$$n_{o2} = n_2 - n_{c2} \qquad (3\text{-}2)$$

$$n_w = n_{c1} + n_{c2} - n_{a1} \qquad (3\text{-}3)$$

For multiple-lane two-way highways,

$$n_{o1} = n_1 - n_{c1} + n_{c2} \qquad (3\text{-}4)$$

$$n_{o2} = n_2 - n_{c2} \qquad (3\text{-}5)$$

$$n_w = n_{c1} - n_{a1} \qquad (3\text{-}6)$$

(2) Operation Characteristics

The following characteristics in terms of work zone operation are considered in our model:

(2.1) The work zone length $\boldsymbol{L_w}$;

(2.2) The work zone schedule.

In this study, we focus on stationary recurring work zones for which construction window establishes the starting and ending time for the construction activity. The lane closure is limited to the time defined in the time window, such as 9-hour nighttime time windows, 48-hour weekend time windows. Therefore, the work zone schedule can be determined by three variables: the work starting time $\boldsymbol{T_s}$, the work duration $\boldsymbol{D_w}$, the number of time windows needed to finish the project $\boldsymbol{N_t}$.

(3) Work time and cost

Four parameters are required to specify to estimate the work time required to complete a work zone and corresponding maintenance cost per zone.

(3.1) The fixed setup cost per zone z_1.

(3.2) The average additional cost required per zone per mile per lane z_2.

(3.3) The fixed setup time per zone z_3.

(3.4) The average additional time required per zone per mile per lane z_4.

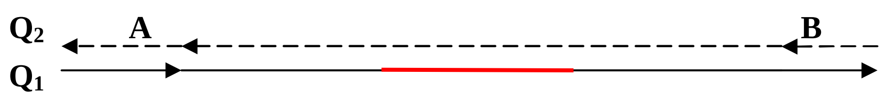
The values of these parameters depend on the maintenance type (patching, grinding, resurfacing, etc.), construction method and lane closure strategies. For example, a crossover may require more time and add extra cost due to installation of more devices. More access lanes may increase the operation efficiency and thereby reduce the work time and cost.

Using the work time parameters, we can assume a linear relation between work zone length L_w and work duration D_w for one zone with the following form:

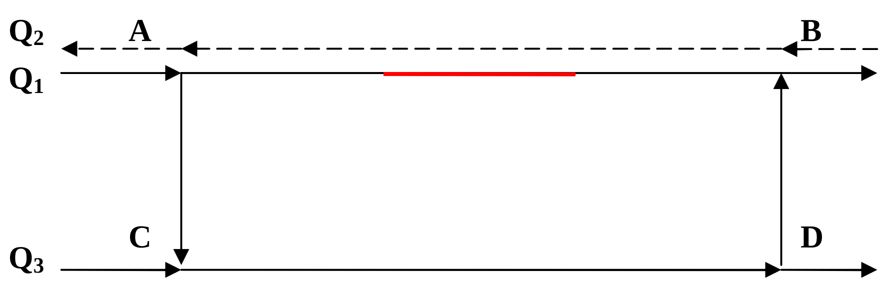
$$D_w = z_3 + n_w\,(z_4)L_w \tag{3-7}$$

(4) Detour Type

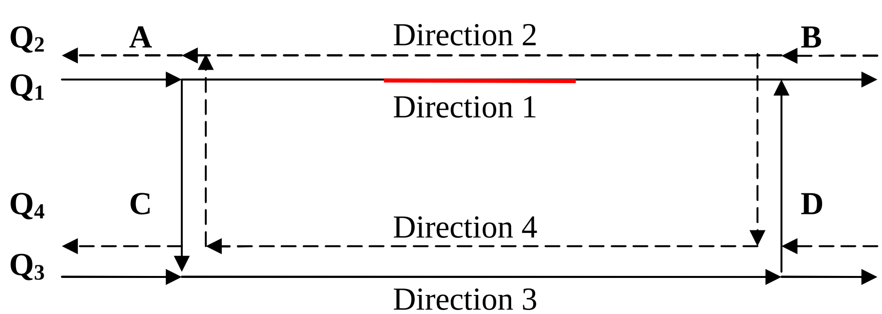
The detour type and corresponding diverted fraction(s) if any detour(s) are available. Several typical detour types are demonstrated in Table 3.1 and Figure 3.3.



(a) Detour Type 1



(a) Detour Type 2



(a) Detour Type 3

Figure 3.3 Roadway Network with Different Detour Types

Table 3.1 Detour Types

Detour Type	Work Zone Link	Diverted Traffic	Detour Route
Type 1	No Detour Available		-
Type 2	Single Detour Available	Q1	AC->CD->DB
Type 3	Single Detour Available for traffic in each direction	Q1 and Q2	AC->CD->DB BD->DC->CA

3.2 Analytic Model for Steady Traffic Inflows

In this subsection, an analytic model to evaluate work zone costs under steady traffic inflows is developed. Since it is a quite simplified model, it is only suitable for roadways with light traffic in a simple network.

3.2.1 Assumptions

Several assumptions made to simplify and formulate this problem are listed below:

1. Traffic demand is steady at all times.
2. A single detour can be considered. For 2-lane 2-way roadways, it is assumed that traffic in both directions can be diverted and detour types 1 and 3 are considered. For multiple-lane 2-way roadways, it is assumed that only the traffic in the direction with work zone may need to be diverted. Thus, detour types 1 and 2 are considered.
3. For 2-lane 2-way highways, no crossover strategy is applied. Two lanes in two ways can be closed simultaneously ($n_{c1}=1$, $n_{c2}=1$). In this case, the traffic in both directions has to be fully diverted to detours ($p_1=1$, $p_2=1$).
4. For 2-lane 2-way highways, queues in both directions will be cleared within each cycle for two-lane two-way highways. Thus, the one-lane work zone capacity must exceed the combined flows of both directions.
5. Traffic moves at a uniform speed through a work zone and at a different uniform speed elsewhere.

6. The effects on speeds of the original detour flows on the relatively short detour segments AC and *DB* in Figure 3.3 are negligible.
7. Possible signal or stop sign delays on the detour AC-CD-DB may be neglected.
8. Queue backups to the maintained road along the first detour *AC* may be neglected.
9. The detour's capacity always exceeds the original flow along the detour.
10. When calculating user delay cost, the value of user time used in numerical analysis is the weighted average cost of driver and passenger's user time for passenger cars and trucks.

3.2.2 Model Formulation

The total cost for one zone is

$$C_T=C_M+C_A+C_D \tag{3-8}$$

where, C_M = Maintenance Cost

C_A = Accident Cost

C_D = Delay Cost

C_D = Total Cost

To compare different work zone characteristics for the same maintenance project, we use the total cost per lane length as the performance measurement. It is calculated as the total cost per zone divided by the product of work zone length and the number of maintained lanes.

$$C_t = \frac{C_T}{L_w \cdot n_w} \tag{3-9}$$

3.2.3.1 Agency Cost

The total agency cost (C_M) for maintaining a zone of length L_w is a linear function with the following form:

$$C_M= z_1+ n_w\,(z_2)L_w \tag{3-10}$$

3.2.3.2 Accident Cost

The accident cost (C_A) incurred by the traffic passing the work zone, can be determined from the number of crashes per 100 million vehicle hours n_a multiplied by the product of the increasing delay (t_d) and the average cost per crash v_a (Chien and Schonfeld, 2001). The accident cost per work zone with length L can be estimated as:

$$C_A = \frac{n_a v_a t_d}{10^8} \tag{3-11}$$

where, n_a = represents the number of accidents per 100 million vehicle hour

v_a = the average cost per accident

t_d = the user delay

3.2.3.3 Delay Cost

The user delay cost per work zone can be obtained as

$$C_D = t_d\, v_d = (t_{d1} + t_{d2} + t_{d3} + t_{d23})\, v_d \tag{3-12}$$

where, v_d = the average user's time value (in \$/veh-hr);

t_d = the total user delay, which is the sum of t_{d1} , t_{d2} and t_{d3};

t_{d1} = the delay along the mainline, which is the sum of moving delay and queuing delay;

t_{d2} = the moving delay of diverted flow if a fraction of flow is diverted;

t_{d3} = the moving delay of original flow on detour if a fraction of flow is diverted;

t_{d23} = the queuing delay upstream the detour route due to the detour's capacity.

(1) User delay along the mainline t_{d1}

The formulation to evaluate user delay along the mainline road is different for two-lane two-way highways and for multiple-lane two-way highways, due to their different work zone characteristics.

(1.1) Two-lane two-way highway

For two-lane two-way highways, the user delay along the mainline consists of the queuing delay upstream of work zones due to a one-way traffic control ($t_{d1,queuing}$) and the moving delay through work zones ($t_{d1,moving}$).

$$t_{d1} = t_{d1,queuing} + t_{d1,moving} \tag{3-13}$$

The queuing delay $t_{d1,queuing}$ per zone is the total delay per control cycle Y in both directions multiplied by the number of cycles N per work zone.

$$t_{d1,queuing} = YN \tag{3-14}$$

where Y = summation of the delays (e.g., Y_1 and Y_2) incurred by the traffic flows from directions 1 and 2 per cycle. Y_1 and Y_2 can be derived by using deterministic queuing analysis.

Schonfeld and Chien (1999) formulated the queuing delay cost per zone along mainline of the work zone area and obtained the following relation:

$$t_{d1,queuing} = \frac{(z_3 + z_4 L_w)[Q_{1m}(C_w - Q_{1m}) + Q_{2m}(C_w - Q_{2m})]}{V_w (C_w - Q_{1m} - Q_{2m})} L_w \qquad (3\text{-}15)$$

where,

C_w = work zone capacity, which is the maximum number of vehicles discharging from the work zone segment;

L_w = work zone length;

V_w = average work zone speed;

Q_{1m} = the traffic flow in direction 1 along work zone link;

$$Q_{1m} = (1-p_1)\, Q_1 \qquad (3\text{-}16a)$$

Q_{2m} = the traffic flow in direction 2 along counter work zone link;

$$Q_{2m} = (1-p_2)\, Q_2 \qquad (3\text{-}16b)$$

p_1 = the diverted fraction for the original traffic flow Q_1 in direction 1;

p_2 = the diverted fraction for the original traffic flow Q_2 in direction 2.

The moving delay cost of the traffic flows Q_{1m} and Q_{2m}, denoted as $t_{d1,moving}$, is the cost increment due to the work zone. It is equal to the flow (Q_{1m} + Q_{2m}) multiplied by: (1) the maintenance duration per zone (z_3+z_4 L_w), and (2) the travel time difference over zone length with the work zone, L/V_W, and without the work zone, L/V_0. Thus:

$$t_{d1,moving} = (Q_{1m} + Q_{2m})(z_3 + z_4 L_w)\left(\frac{L_w}{V_w} - \frac{L_w}{V_0}\right) \qquad (3\text{-}17)$$

where V_0 represents the speed on the mainline without any work zone.

Therefore, the user delay along the mainline for the two-lane two-way highways can be obtained for the following formulation:

$$t_{d1} = t_{d1,queuing} + t_{d1,moving}$$

$$= \frac{(z_3 + z_4 L_w)[Q_{1m}(C_w - Q_{1m}) + Q_{2m}(C_w - Q_{2m})]}{V_w (C_w - Q_{1m} - Q_{2m})} L_w + (Q_{1m} + Q_{2m})(z_3 + z_4 L_w)\left(\frac{L_w}{V_w} - \frac{L_w}{V_0}\right) \qquad (3\text{-}18)$$

(1.2) Multiple-lane two-way highway

For multiple-lane two-way highways,

(1) If no crossover strategy is applied, the traffic flow in direction 2 is not affected by work zone activities. Thus the user delay along the mainline (t_{d1}) is the user delay in direction 1 (t_{d11}) including the queuing delay and the moving delay;

(2) If crossover strategy is applied, the traffic flows in both directions is affected by work zone activities. The user delay along the mainline (t_{d1}) is the sum of the user delay in direction 1 (t_{d11}) and ay in direction 2 (t_{d12}). The queuing delay and the moving delay of traffic flows in both directions should be calculated.

$$t_{d1} = t_{d11} + \xi_c t_{d12}$$
$$= (t_{d11,queuing} + t_{d11,moving}) + \xi_c (t_{d12,queuing} + t_{d12,moving}) \quad (3\text{-}19)$$

where t_{d1} = the delay along the mainline, which is the sum of t_{d11} and t_{d12}

t_{d11} = the delay for the traffic flow Q_{1m} from direction 1 along work zone link.

t_{d12} = the delay for the traffic flow Q_{2m} from direction 2 along counter work zone link.

ξ_c = 1 if crossover strategy is applied, otherwise ξ_c=0

The following variables are defined:

Q_{1m} = approaching traffic flow in Direction 1 along the mainline, $Q_{1m}= (1-p_1) Q_1$

Q_{2m} = approaching traffic flow in Direction 2 along the mainline, $Q_{2m}= (1-p_2) Q_2$

c_w = work zone capacity per lane (veh/hr per lane)

c_0 = normal road capacity in normal per lane (veh/hr per lane)

c_{w1} = work zone capacity in Direction 1 (veh/hr), $c_{w1}=c_w n_{o1}$

c_{w2} = work zone capacity in Direction 2 if crossover is applied (veh/hr), $c_{w2}=c_w n_{o2}$

c_{01} = normal road capacity in Direction 1 (veh/hr), $c_{01}=c_0 n_1$

c_{02} = normal road capacity in Direction 2 (veh/hr), $c_{02}=c_0 n_2$

D_w = maintenance duration per zone, $D_w=z_3+z_4L_w$

If Q_{im} exceeds the work zone capacity c_{wi}, a queue forms, which then dissipates when the closed lane is open again, as shown in Figure 3.4. Based on the deterministic queuing model, the queuing delay per zone can be obtained from the following equations:

$$t_{d1i,queuing} = 0 \qquad \text{when } \mathrm{Q_{mi}} \le \mathrm{c_{wi}} \; i=1,2 \qquad (3\text{-}20a)$$

$$t_{d1i,queuing} = \frac{1}{2}\left(1+\frac{Q_{mi}-c_{wi}}{c_{0i}-Q_{mi}}\right)(Q_{mi}-c_{wi})(z_3+z_4L_w)^2 \qquad \text{when } \mathrm{Q_{mi}} > \mathrm{c_{wi}} \; i=1,2 \qquad (3\text{-}20b)$$

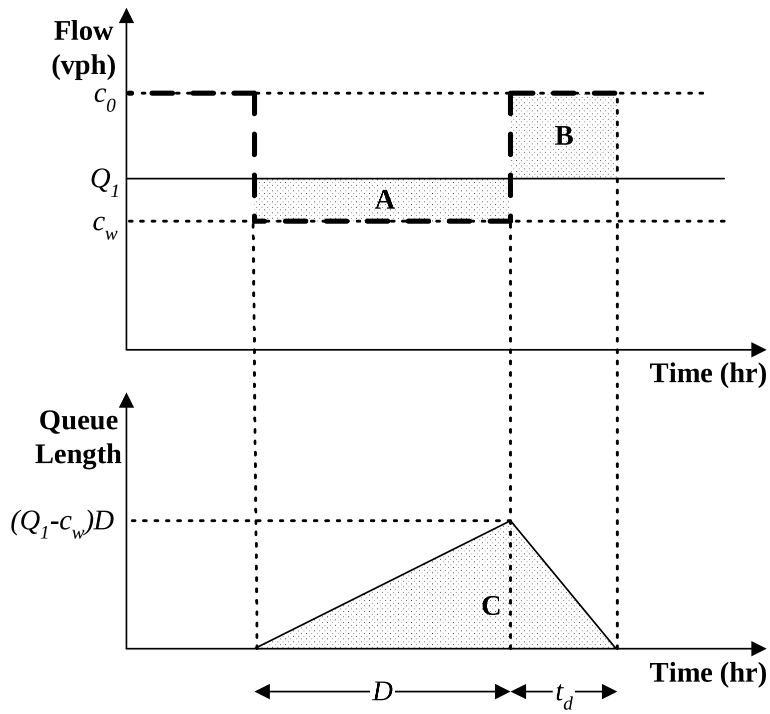


Figure 3.4 Deterministic Queuing Model (from Chien and Schonfeld, 2001)

The moving delay $t_{d1i,moving}$ is a function of the difference between the travel time on a road with and without a work zone. . If Q_{mi} is less than c_{wi}, the flow passing through the work zone is Q_{mi}. If Q_{mi} is more than c_{wi}, the maximum flow allowed to pass through the work zone is c_{wi}. Then the moving delays are determined with equations (3-21a) and (3-21b):

$$t_{d1i,moving} = (\frac{L_w}{V_w} - \frac{L_w}{V_{AB}})Q_{mi}(z_3 + z_4 L_w) \qquad \text{when } Q_{mi} \le c_{wi} \tag{3-21a}$$

$$t_{d1i,moving} = (\frac{L_w}{V_w} - \frac{L_w}{V_{AB}})c_{wi}(z_3 + z_4 L_w) \qquad \text{when } Q_{mi} > c_{wi} \tag{3-21b}$$

(2) Moving delay of the diverted flow

When a detour is used, the user delay of the traffic flow which is diverted from the mainline to the detour should be considered as a part of the total user delay.

The user moving delay cost of the diverted flow p_1Q_1 from Direction 1, denoted as t_{d2}, is equal to the flow $Q_{1d} = pQ_1$ multiplied by: (1) the average maintenance duration per zone (z_3+z_4 L_w) , and (2) the time difference between the time vehicles through the detour, $\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V_{CD}}$, and the time vehicles through the maintained road AB without work zone,

$\frac{L_{AB}}{V_{AB}}$. If p_2Q_2 is also diverted from Direction 2, we do the same thing and sum up the two parts. Thus:

$$t_{d2} = \xi_1 p_1 Q_1 (z_3 + z_4 L_w)(\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{AB}}{V_{AB}}) + \xi_2 p_2 Q_2 (z_3 + z_4 L_w)(\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,DC}} - \frac{L_{AB}}{V_{AB}}) \quad (3\text{-}22)$$

L_{AC}, L_{BD}, L_{CD} are the lengths of the segments of the detour shown in Figure 3.3. V_{AB} represents the speed on the maintained road without any work zone while *VBD* and VAC are the average speed on the detour road.

V^*_{CD} and V^*_{DC} are the detour speeds affected by diverted traffic in Direction 3 and 4. Both speeds can be derived from speed-density relation basing on the following procedures.

ξ_1 = 1 if the traffic in Direction 1 will be diverted to the Direction 3, otherwise ξ_1=0;

ξ_1 = 1 if the traffic in Direction 2 will be diverted to the Direction 4, otherwise ξ_1=0.

In traffic flow theory, the relation among flow Q, density K, and speed V is:

$$Q = KV \quad (3\text{-}23)$$

The speed function can be formulated by applying Greenshield's model (Gerlough and Huber, 1975), which is widely used in practice:

$$V = V_f - \frac{V_f}{K_j} K \quad (3\text{-}24)$$

where V_f is free flow speed, K_j is jam density.

Substituting Eq. (3-24) into (3-23), we obtain

$$Q = K_j V - \frac{K_j}{V_f} V^2 \quad (3\text{-}25)$$

Given the Vf and Kj, we can derive the detour capacity:

$$C=K_j V_f /4 \quad (3\text{-}26)$$

If the detour has enough capacity for traffic inflows, the traffic inflow along the detour is the sum of the original flow and the diverted flow. If the traffic inflow exceeds the detour capacity, queues will form upstream the detour road and the maximum flow allowed to pass through the detour is C. We can calculate the speeds on the detour ($V^*_{d,CD}, V^*_{d,DC}$) by solving the quadratic relation (3-24) with the following equations.

$$V^*_{d,CD} = \frac{K_j V_{CD} + \sqrt{(K_j V_{CD})^2 - 4K_j V_{CD}(Q_3 + p_1 Q_1)}}{2K_j} \quad \text{if } C_{CD} > Q_3 + P_1 Q_1 \quad (3\text{-}27a)$$

$$V_{d,CD}^{*} = \frac{V_{CD}}{2} \qquad \text{if } C_{CD} \leqslant Q_3 + P_1 Q_1 \qquad (3\text{-}27b)$$

$$V_{d,DC}^{*} = \frac{K_j V_{DC} + \sqrt{(K_j V_{DC})^2 - 4K_j V_{DC}(Q_4 + p_2 Q_2)}}{2K_j} \qquad \text{if } C_{DC} > Q_4 + P_2 Q_2 \qquad (3\text{-}28a)$$

$$V_{d,CD}^{*} = \frac{V_{CD}}{2} \qquad \text{if } C_{DC} \leqslant Q_4 + P_2 Q_2 \qquad (3\text{-}28b)$$

(3) Moving delay of the original flow on detour

When a detour is used, the delay of the original flow on the detour, as affected by the diverted flow, should also be considered. Denoted as t_{d3}, it equals the flow multiplied by: (1) the average maintenance duration per kilometer and (2) the travel time difference over L_{CD} between with the diverted flow and without it.

$$t_{d3} = \xi_1 Q_3 (z_3 + z_4 L_w)\left(\frac{L_{CD}}{V_{d,CD}^{*}} - \frac{L_{CD}}{V_{CD}}\right) + \xi_2 Q_4 (z_3 + z_4 L_w)\left(\frac{L_{CD}}{V_{d,DC}^{*}} - \frac{L_{CD}}{V_{DC}}\right) \qquad (3\text{-}29)$$

where V_{CD} and V_{DC} represent the original speeds on the detour unaffected by the diverted flows.

(4) Queuing delay upstream the detour

When a detour is used, queuing delay upstream the detour should be calculated when the sum of original flow on detour and the diverted flow exceeds the detour's capacity. Based on the deterministic queuing model, the queuing delay upstream the detour for one wok zone can be obtained from the following equations:

$$t_{d23} = \xi_1 t_{d23,CD} + \xi_2 t_{d23,DC} \qquad (3\text{-}30)$$

where $t_{d23,CD}$ and $t_{d23,DC}$ represent the queuing delay upstream detour in Direction 3 and Direction 4:

$$t_{d23,CD} = 0 \qquad \text{when } Q_3 + P_1 Q_1 \le C_{CD} \qquad (3\text{-}31a)$$

$$t_{d23,CD} = \frac{1}{2}\left(1 + \frac{Q_3 + P_1 Q_1 - C_{CD}}{C_{CD} - Q_3}\right)(Q_3 + P_1 Q_1 - C_{CD})(z_3 + z_4 L_w)^2 \qquad \text{when } Q_3 + P_1 Q_1 > C_{CD} \qquad (3\text{-}31b)$$

$$t_{d23,DC} = 0 \qquad \text{when } Q_4 + P_2 Q_2 \le C_{DC} \qquad (3\text{-}32a)$$

$$t_{d23,DC} = \frac{1}{2}\left(1 + \frac{Q_4 + P_2 Q_2 - C_{DC}}{C_{DC} - Q_4}\right)(Q_4 + P_2 Q_2 - C_{DC})(z_3 + z_4 L_w)^2 \qquad \text{when } Q_4 + P_2 Q_2 > C_{DC} \qquad (3\text{-}32b)$$

3.3 Analytic Model for Time-Dependent Traffic Inflows

For steady traffic inflow cases, only the closure durations affect the costs while the specific schedule (e.g. starting and stopping times) makes no difference. However, when traffic inflows vary over time, the lane closure schedules will significantly affect the costs.

In this subsection, the analytic work zone cost model considering time-dependent traffic inflows are presented.

3.3.1 Assumptions

Several assumptions made to simplify and formulate this problem are listed below:

1. Traffic demand varies over time. An hour is used as a duration unit in which traffic inflows stay appropriately constant.
2. A single detour can be considered. For 2-lane 2-way roadways, it is assumed that traffic in both directions can be diverted and detour types 1 and 3 are considered. For multiple-lane 2-way roadways, it is assumed that only the traffic in the direction with work zone may need to be diverted. Thus, detour types 1 and 2 are considered.
3. For 2-lane 2-way highways, no crossover strategy is applicable. Two lanes in two ways can be closed simultaneously ($n_{c1}=1, n_{c2}=1$). In this case, the traffic in both directions must be fully diverted to detours ($p_1=1, p_2=1$).
4. For 2-lane 2-way highways, queues in both directions are cleared within each cycle for two-lane two-way highways. Thus, the one-lane work zone capacity always exceeds the combined flows of both directions.
5. The effects on speeds of the original detour flows on the relatively short detour segments AC and *DB* in Figure 3.3 are negligible.
6. Possible signal or stop sign delays on the detour can be considered.
7. Queue backups to the maintained road may be neglected.
8. The roadway capacity always exceeds its original flow under the normal situation without a work zone.
9. When calculating user delay cost, the value of user time used in numerical analysis is the weighted average cost of driver and passenger's user time for passenger cars and trucks.
10. User delays including queuing and moving delays are estimated with deterministic models. Delays caused by acceleration, deceleration and shock-waves are neglected.

11. Traffic speeds along the detour route CD are estimated with a deterministic traffic flow model (Greenshield's model).
12. The detour's capacity always exceeds its original flows.
13. The user delay cost is represented by a constant average cost per vehicle hour v_d.
14. The diverted fraction does not vary for the duration of a work zone if a detour is used.
15. The same construction time window and work zone configurations are repeated throughout the whole project. No mixed construction windows are considered.

3.3.2 Model Formulation

As in the analytic model for steady traffic inflows, the total cost per maintained zone length is used as the performance measurement. It is calculated as the sum of maintenance cost, accident cost and user delay cost multiplied by the production of work zone length and the maintained lanes. The total cost per lane length is

$$C_t = \frac{C_T}{L_w \cdot n_w} = \frac{C_M + C_A + C_D}{L_w \cdot n_w} \tag{3-33}$$

where C_t = the total cost per lane length

C_T = the total cost per zone

C_M = Maintenance cost per zone;

C_A = Accident cost per zone

C_D = Delay cost per zone

L_w = the work zone length

n_w = the number of maintained lanes in a work zone

3.3.2.1 Maintenance Cost

The maintenance cost (C_M) for maintaining a zone of length L_w is a linear function with the following form:

$$C_M = z_1 + (z_2)\, n_w L_w \tag{3-34}$$

3.3.2.2 Accident Cost

The crash cost (C_A) incurred by the traffic passing the work zone is estimated as:

$$C_A = \frac{\theta_a n_a v_a t_d}{10^8} \tag{3-35}$$

where n_a = the number of accidents per 100 million vehicle hours

v_a = the average cost per accident

t_d = the user delay

3.3.2.3 Delay Cost

The user delay cost per work zone can be obtained as:

$$C_D = t_d\, v_d = (t_{d1} + t_{d2} + t_{d3} + t_{d23})\, v_d \qquad (3\text{-}36)$$

Where v_d = the average user's time value (in \$/veh-hr);

t_d = the total user delay, which is the sum of t_{d1}, t_{d2} and t_{d3};

t_{d1} = the delay along the mainline, which is the sum of moving delay and queuing delay;

t_{d2} = the moving delay of diverted flow if a fraction of flow is diverted to a detour route;

t_{d3} = the moving delay of original flow on detour if a fraction of flow is diverted to a detour route;

t_{d23} = the queuing delay upstream the detour route due to the detour's capacity.

For time-dependent traffic inflows, assume that work zone is maintained over n duration units and D_i (i =1, 2, …., n) represents a duration unit in which inflows stay appropriately constant. Then the duration for the work zone is

$$D_w = \sum_{i=1}^{n} D_i \qquad (3\text{-}37)$$

The work zone length can be derived from the work zone duration

$$L_w = \frac{D_w - z_3}{n_w z_4} \qquad (3\text{-}38)$$

(1) User delay along the mainline t_{d1}

The user delay along the mainline (t_{d1}) includes the queuing delay occurring before the work zone ($t_{d1,queuing}$) and the moving delay experienced by drivers traveling through the work zone ($t_{d1,moving}$).

$$t_{d1} = t_{d1,queuing} + t_{d1,moving} \qquad (3\text{-}39)$$

(1.1) Two-lane two-way highway

Based on the queuing delay model developed by Schonfeld and Chien (2001) for two-lane two-way highway with time-dependent traffic inflows, the following equation is formulated to calculate the queuing delay caused by time-dependent one-way traffic control on two-lane two-way highways with one lane closed.

$$t_{d1,queuing} = \sum_{i}^{n} \frac{[Q_{1m}^{i}(C_w - Q_{1m}^{i}) + Q_{2m}^{i}(C_w - Q_{2m}^{i})]}{V_w(C_w - Q_{1m}^{i} - Q_{2m}^{i})} D_i L_w \tag{3-40}$$

Where C_w = work zone capacity, which is the maximum number of vehicles discharging from the work zone segment;

L_w = the length of work zone;

V_w = average work zone speed;

Q^i_{1m} = the traffic flow in direction 1 in duration unit i;

$$Q^i_{1m} = (1-p_1)\, Q^i_1 \tag{3-41a}$$

Q^i_{2m} = the traffic flow in direction 2 in duration unit i;

$$Q^i_{2m} = (1-p_2)\, Q^i_2 \tag{3-41b}$$

p_1 = the diverted fraction for the original traffic flow Q^i_1 in direction 1;

p_2 = the diverted fraction for the original traffic flow Q^i_2 in direction 2.

The moving delay for a work zone in each period D_i is equal to the flow ($Q^i_{1m} + Q^i_{2m}$) multiplied by: (1) the period, D_i, (2) the travel time difference over the zone length L_w with the work zone, $\frac{L_w}{V_w}$, and without the work zone, $\frac{L_w}{V_0}$. Thus:

$$t_{d1,moving} = \sum_{i}^{n} (Q_{1m}^{i} + Q_{2m}^{i}) D_i \left(\frac{L_w}{V_w} - \frac{L_w}{V_0}\right) \tag{3-42}$$

(1.2) Multiple-lane two-way highway

Under time-dependent traffic flow, queuing delay ($t_{d1,queuing}$) for multiple-lane freeway work zone are computed numerically using deterministic queuing model. The method is illustrated in Figure 3.5.

Within duration D_i, if the inflow along mainline Q^i_{1m} exceeds the work zone capacity c_w, a queue forms. The cumulative number of vehicles in a queue within D_i is

$$q_{\mathrm{i}} = q_{i-1} + (Q_{1m}^{i} - c_w) D_i \qquad \text{when } Q_{1m}^{i} \geq c_w \tag{3-43}$$

If the inflow Q^i_{1m} does not exceed the work zone capacity c_w, the queuing delay time is zero and the existing queue starts to dissipate. In this case, the cumulative number of vehicles in a queue within D_i is

$$q_{\text{i}} = \max\{q_{i-1} - s_i, 0\} \qquad \text{when } Q_i < c_w \tag{3-44}$$

where s_i represents the queue reduction within D_i.

$$s_i = c_w - Q_{1m}^i \tag{3-45}$$

If the work zone duration ends before D_i, the roadway recovers its original capacity c_0.

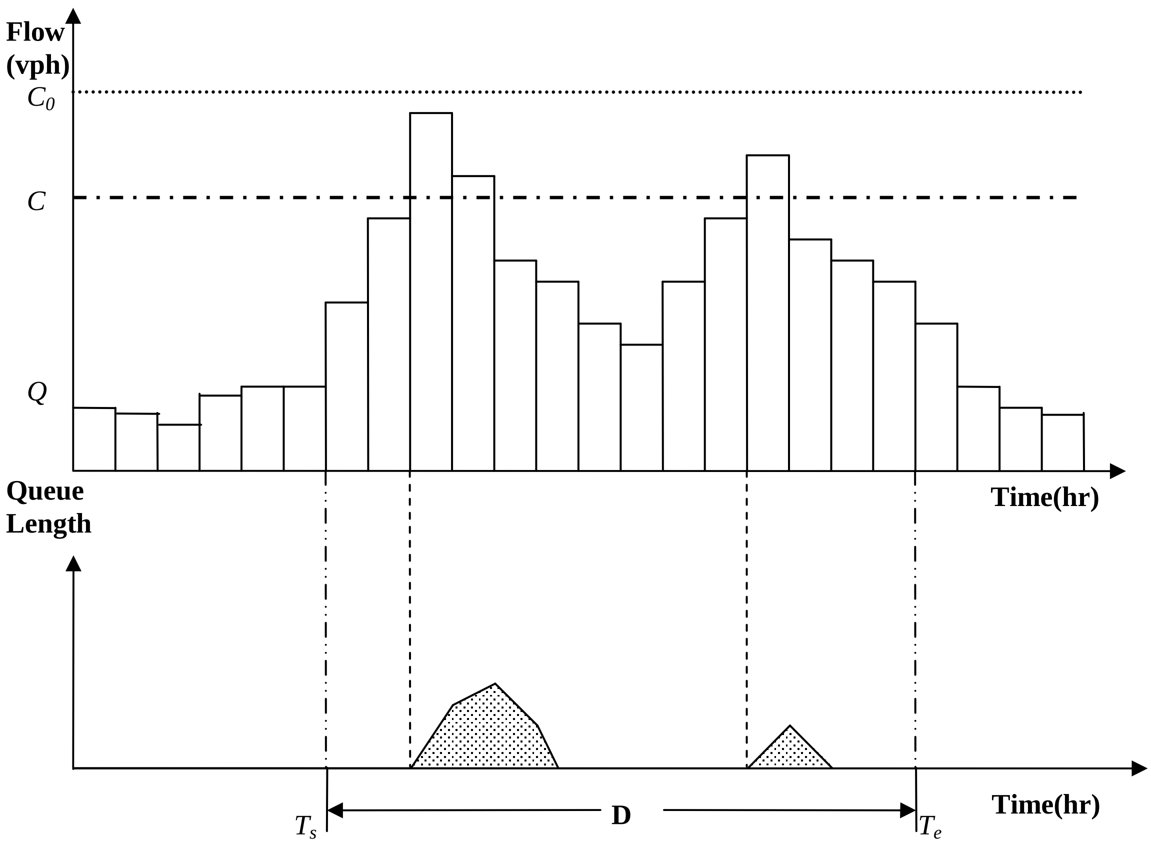
$$q_{\text{i}} = \max\{q_{i-1} - s_i, 0\} \qquad \text{when } Q_i < c_0 \tag{3-46}$$

where

$$s_i = c_0 - Q_{1m}^i \tag{3-47}$$

The total queuing delay ($t_{d1,queuing}$) is obtained as

$$t_{d1,queuing} = \sum_{i=1}^{n} \frac{q_i + q_{i+1}}{2} D_i \tag{3-48}$$



Q: Hourly Traffic Flow Approaching the Work Zone (vph)
C_0: Roadway Capacity without Work Zone (vph)
C_w: Roadway Capacity with Work Zone (vph)
D: Constrained Construction Time Window
T_s: Earliest Work Zone Starting Time
T_e: Latest Work Zone Ending Time
: Queuing Delay (veh.hr)

Figure 3.5 Queuing Delay Estimated by Deterministic Queuing Model

The moving delay in a duration unit ($t^{i}_{d1,moving}$) is obtained by multiplying the outflow passing through the work zone by the difference between the travel time on a road with and without a work zone.

Within duration D_i, if the total volume, which is the sum of inflow Q^{i}_{1m} and queue length accumulated from the previous duration q_{i-1}, can be discharged in D_i, the moving delay $t^{i}_{d1,moving}$ can be obtained from Eq. (3-49a). If not, the moving delay $t^{i}_{d1,moving}$ is calculated from Eq.(3-49b) .

$$t^{i}_{d1,moving} = (\frac{L_{AB}}{V_w} - \frac{L_{AB}}{V_0})(Q^{i}_{1m} + q_{i-1})D_i \qquad \text{when } (Q^{i}_{1m} + q_{i-1}) \le \mathrm{c_w} D_i \qquad \text{(3-49a)}$$

$$t^{i}_{d1,moving} = (\frac{L_{AB}}{V_w} - \frac{L_{AB}}{V_0})c_w D_i \qquad \text{when } (Q^{i}_{1m} + q_{i-1}) > \mathrm{c_w} D_i \qquad \text{(3-49b)}$$

Therefore, the total moving delay during the work zone duration is

$$t_{d1,moving} = \sum_{i=1}^{n} t^{i}_{d1,moving} \qquad \text{(3-50)}$$

where V_0 = average approaching speed;
V_w = average work zone speed.

Based on the assumption that the capacity without work zone is enough for the entry inflow, free-flow speed can be used to approximate the average approaching speed V_0. The average work zone speed can be the speed limit along the work zone.

(2) Moving delay of the diverted flow

When a detour is used, the moving delay of the traffic flow which is diverted from the mainline to the detour can be calculated by the following equations:

$$t_{d2} = \xi_1 \sum_{i=1}^{n} p_1 Q^{i}_{1} D_i (\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^{*,i}_{d,CD}} + T_{\text{int}} - \frac{L_{AB}}{V_{AB}})$$

$$+ \ \xi_2 \sum_{i=1}^{n} p_2 Q^{i}_{2} D_i (\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^{*,i}_{d,DC}} + T_{\text{int}} - \frac{L_{AB}}{V_{AB}}) \qquad \text{(3-51)}$$

L_{AC}, L_{BD}, L_{CD} are the lengths of the segments of the detour shown in Figure 3.3. V_{AB} represents the speed on the maintained road without any work zone while V_{BD} and V_{AC} are the average speed on the detour road.

T_{int} represents the average waiting time passing intersections along the detour route.

$V_{d,CD}^{*,i}$ and $V_{d,DC}^{*,i}$ are the detour speeds in period D_i affected by diverted traffic in Direction 3 and 4 shown in Figure 3.3. Both speeds can be derived from speed-density relation defined in Greenshield's model.

$$V_{d,CD}^{*,i} = \frac{K_j V_{CD} + \sqrt{(K_j V_{CD})^2 - 4K_j V_{CD}(Q_3^i + p_1 Q_1^i)}}{2K_j} \quad if\ C_{CD} > Q^i_3 + P_1 Q^i_1 \qquad \text{(3-52a)}$$

$$V_{d,CD}^{*,i} = \frac{V_{CD}}{2} \quad if\ C_{CD} \leqslant Q^i_3 + P_1 Q^i_1 \qquad \text{(3-52b)}$$

$$V_{d,DC}^{*,i} = \frac{K_j V_{DC} + \sqrt{(K_j V_{DC})^2 - 4K_j V_{DC}(Q_4^i + p_2 Q_2^i)}}{2K_j} \quad if\ C_{DC} > Q^i_4 + P_2 Q^i_2 \qquad \text{(3-53a)}$$

$$V_{d,CD}^{*,i} = \frac{V_{CD}}{2} \quad if\ C_{DC} \leqslant Q^i_4 + P_2 Q^i_2 \qquad \text{(3-53b)}$$

ξ_1 = 1 if the traffic in Direction 1 will be diverted to the Direction 3, otherwise ξ_1=0;

ξ_2 = 1 if the traffic in Direction 2 will be diverted to the Direction 4, otherwise ξ_2=0.

(3) Moving delay of the original flow on detour

When detour strategy is applied, the delay of the original flow on the detour, as affected by the diverted flow, can be obtained from the following equation:

$$t_{d3} = \xi_1 \sum_{i=1}^{n} Q_3^i D_i \left(\frac{L_{CD}}{V_{d,CD}^{*,i}} - \frac{L_{CD}}{V_{CD}}\right) + \xi_2 \sum_{i=1}^{n} Q_4^i D_i \left(\frac{L_{CD}}{V_{d,DC}^{*,i}} - \frac{L_{CD}}{V_{DC}}\right) \qquad \text{(3-54)}$$

(4) Queuing delay upstream the detour

When a detour is used, queuing delay upstream of the detour ($t_{d23,queuing}$) should be considered. Similarly to the estimation of queuing delay upstream the mainline road, queuing delay upstream the detour road for time-dependent traffic inflows can be computed numerically using a deterministic queuing model.

Within duration D_i, if the inflow along the mainline ($Q_3^i + p_1 Q_1^i$) exceeds the detour capacity C_{CD}, a queue forms. The cumulative number of vehicles in a queue within D_i is

$$q_i = q_{i-1} + (Q_3^i + p_1 Q_1^i - c_w) D_i \qquad \text{when } Q_3^i + p_1 Q_1^i \geq C_{CD} \qquad (3\text{-}55)$$

If the inflow ($Q_3^i + p_1 Q_1^i$) does not exceed the work zone capacity c_w, the queuing delay time is zero and the existing queue starts to dissipate. In this case, the cumulative number of vehicles in a queue within D_i is

$$q_i = \max\{q_{i-1} - s_i, 0\} \qquad \text{when } Q_3^i + p_1 Q_1^i < C_{CD} \qquad (3\text{-}56)$$

where s_i represents the queue reduction within D_i.

$$s_i = C_{CD} - (Q_3^i + p_1 Q_1^i) \qquad (3\text{-}57)$$

If the work zone duration ends before D_i, the roadway recovers its original capacity c_0.

$$q_i = \max\{q_{i-1} - s_i, 0\} \qquad \text{when } Q_3^i < c_0 \qquad (3\text{-}58)$$

where $s_i = C_{CD} - Q_3^i$ (3-59)

The total queuing delay in direction 3 ($t^3_{d23,queuing}$) is obtained as

$$t^3_{d23,queuing} = \sum_{i=1}^{n} \frac{q_i + q_{i+1}}{2} D_i \qquad (3\text{-}60)$$

If a fraction of the traffic flow in Direction 2 is diverted to the detour in Direction 4 (Figure 3.3), the queuing delay in Direction 4 ($t^4_{d23,queuing}$) should be included using the same procedure as above.

The total queuing delay upstream of the detour is obtained as:

$$t_{d23,queuing} = \xi_1 (t^3_{d23,queuing}) + \xi_2 (t^4_{d23,queuing}) \qquad (3\text{-}61)$$

3.4 Numerical Examples

3.4.1 Problem Description

In this numerical example, a work zone on a four-lane two-way highway considering single detour is analyzed. The performance comparison of various work zone characteristics is examined in this section. The data for this example are summarized in Table 3.2. Table 3.3 shows the AADT and hourly traffic distributions on the maintained road and the detour.

Table 3.2 Notation and Baseline Numerical Inputs

Variable	Description	Value
L_{AB}	Length of Segment AB	3.11 mile
L_{AC}	Length of Segment AC	0.93 mile
L_{CD}	Length of Segment CD	0.93 mile
L_{DB}	Length of Segment DB	2.49 mile
NAB	Number of lanes in Segment AB	2
NCD	Number of lanes in Segment CD	1
Kj	Jam Density	200 veh/mile
c_0	Maximum discharge rate without work zone	2,000 vph /lane
c_w	Maximum discharge rate with work zone	1,300 vph /lane
V_{AB}	Average approaching speed	65 mph
Vw	Average work zone speed	35 mph
V_{CD}	Free flow speed in Segment CD	45 mph
$V_{AC/DB}$	Average speed in Segment AC/DB	45 mph
T_{int}	Average waiting time passing intersections along the detour	30 seconds/veh
n_a	Number of crashes per 100 million vehicle hours	40 acc/100mvh
z_1	Fixed setup cost	1,000 $/zone
z_2	Average maintenance cost per lane kilometer	33,000 $/lane.mile
z_3	Fixed setup time	2 hr/zone
z_4	Average maintenance time per lane kilometer	10 hr/lane.mile
v_a	Average accident cost	142,000 $/accident
v_d	Value of user time	12 $/veh·hr

Table 3.3 AADT and Hourly Traffic Distribution on a Four-Lane Two-Way Freeway

Time Period	Time	Q1	Q2	Q3
0	0:00-1:00	220	930	392
1	1:00-2:00	157	645	391
2	2:00-3:00	148	301	367
3	3:00-4:00	198	238	432
4	4:00-5:00	448	240	432
5	5:00-6:00	1425	326	432
6	6:00-7:00	2941	580	734
7	7:00-8:00	3541	887	1276
8	8:00-9:00	2897	977	1505
9	9:00-10:00	2509	1134	1363
10	10:00-11:00	1793	1283	951
11	11:00-12:00	1586	1589	772
12	12:00-13:00	1528	1544	700
13	13:00-14:00	1475	1673	670
14	14:00-15:00	1541	2074	773
15	15:00-16:00	1414	2808	954
16	16:00-17:00	1079	3501	1042
17	17:00-18:00	957	3719	1026
18	18:00-19:00	991	3061	832
19	19:00-20:00	779	2171	770
20	20:00-21:00	554	1433	644
21	21:00-22:00	504	1314	559
22	22:00-23:00	436	905	392
23	23:00-24:00	325	720	391
AADT		29446	34053	17800
Average Hourly Volume		1227	1418	742

3.4.2 Sensitivity Analysis

Suppose that one lane is closed in the work zone area and no fraction of mainline volume is diverted. We intend to examine the impact of work zone length on the work zone costs in work zone cost model with steady traffic flow and the impact of work scheduling on the costs in the work zone cost model with time-dependent flow. When using the cost model for steady traffic inflows, we use the average hourly volumes for Q_1, Q_2 and Q_3 as the steady traffic inflows.

(1) The analytic model for steady traffic inflows

Chien et al (2001, 2002) proposed that longer zones tend to increase the user delays, but the maintenance activities can be performed more efficiently with fewer repeated setups in longer zones. From Figure 3.6, we can see that, with increasing work zone lengths, the average agency cost decrease while the average user cost increase. Due to the tradeoffs between maintenance cost and user cost, the total cost is a convex function of work zone length and an optimal work zone length with minimum cost can be easily found.

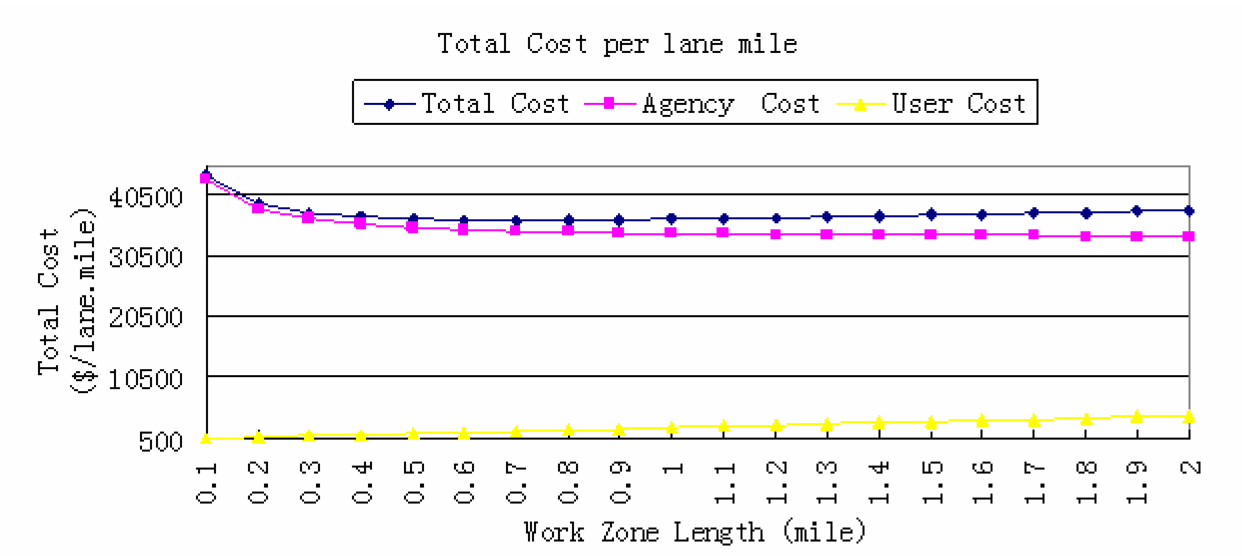


Figure 3.6 Work Zone Cost vs. Work Zone Length
(Cost Model for Steady Traffic Inflows)

(2) The analytical model for time-dependent traffic inflows

For time-dependent traffic inflows, the work start time as well as the duration and the length is important for estimating work zone costs because the influence of work zone activities on the traffic depends on the timing of lane closures. Figure 3.7 shows the work zone costs with varying work zone start times and work durations. As expected, the costs are lower when the work zone takes place during off-peak periods. We can see that the cost function is not convex and multiple local minimum costs may occur.

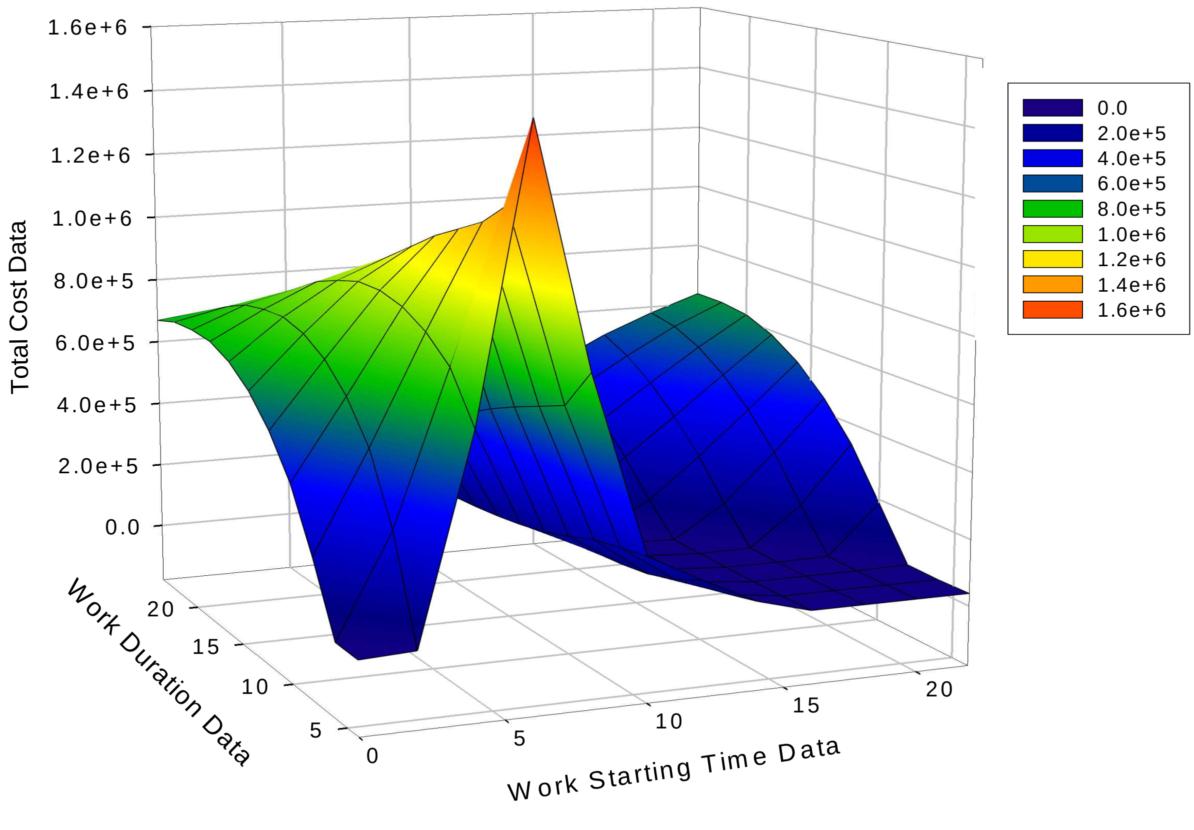


Figure 3.7 Work Zone Cost vs. Work Start time and Work Duration (Cost Model for Time-Dependent Traffic Inflows)

3.4.3 Summary

In this chapter, various highway systems studied in our research and the significant work zone characteristics are first defined. After that, a work zone cost model, which can evaluate the agency cost, the user delay cost and the accident cost based on given work zone characteristics, is developed. Analytic models are developed to estimate user delay costs under steady traffic inflows and time-dependent traffic inflows for two-lane highway and multiple-lane highway work zones. Numerical examples are used to show the impacts of work zone length, work zone duration and work start time on the total cost for steady and time-dependent traffic inflows. The results can provide useful information for developing work zone optimization methods in later chapters.

Chapter 4 Work Zone Simulation Model

The accuracy of user delay estimates significantly affects the measure of the total work zone cost. Microscopic simulation programs, which model each vehicle as a separate entity, are usually expected to provide more accurate estimates of vehicle speeds and delays compared to analytical procedures, especially when the traffic conditions or roadway networks are complex.

In this chapter, CORSIM (Corridor Simulator), a comprehensive microscopic traffic simulation model developed by the Federal Highway Administration (FHWA), is used to simulate various work zone conditions and estimate the user delay and vehicle operating cost due to capacity reductions in freeway work zones.

4.1 Introduction to CORSIM

Traffic simulation models can be classified into microscopic, macroscopic or mesoscopic. Microscopic models address and describe the movement of each individual vehicle in the traffic flow independence of the movement of the adjacent vehicles, both in the longitudinal (car-following behaviour) and in the lateral (lane-changing behaviour) sense. Macroscopic models describe the traffic flow as a fluid with particular characteristics via the aggregate traffic variables traffic density, flow, and mean speed. Mesoscopic models track individual vehicles but group them into platoons with same behaviors, and thus provide the precise level in the middle of microscopic and macroscopic simulation models.

CORSIM is a microscopic and stochastic simulator. It represents single vehicles entering the road network at random times moving second-by-second according to local interaction rules that describe governing phenomena such as car following logic, lane changing, response to traffic control devices, and turning at intersections according to prescribed probabilities.

CORSIM combines two of the most widely used traffic simulation models, NETSIM for surface streets, and FRESIM for freeways. CORSIM simulates traffic and traffic control systems using commonly accepted vehicle and driver behavior models and it has great ability to model complex road networks, various traffic conditions and different traffic control alternatives. CORSIM can handle networks of up to 500 nodes and 1,000 links containing up to 20,000 vehicles at one time.

For matching field observations and predicting traffic performance correctly, the CORSIM simulation model should be properly calibrated for field conditions.

4.2 Simulating Work Zone Conditions in CORSIM

For the CORSIM model, the input data specified by the user consists of a sequence of "record types", which contains a specific set of data items as well as an identification number. These data specified in a "record type" are called "entries".

In NETSIM and FRESIM, different record types are used to contain work zone related information.

4.2.1 NETSIM

(1) Record Type 11

The record type 11 is the NETSIM link description, which describes the geometry and the traffic characteristics of NETSIM links.

The entries 11-17 specify the channelization for all defined lanes. We can simulate a link with one or more closed lanes, by setting proper values of the channelization codes for corresponding lanes. A closed lane can be treated as a transient condition that is due to a construction zone. The entries 23 and 24 specify the mean startup delay and the mean queue discharge headway (in tenths of a second), which may affect signalized intersection capacity in a NETSIM link.

However, only full lanes can be channelized and the capacity of the whole link can be changed. If we want to simulate a work zone segment within a surface street link, we have to divide the link into several links. Also, the road capacity along work zone segment and drivers' behavior characteristics are still hard to calibrate in NETSIM. Therefore, in this project, we will focus on freeway work zones.

(2) Record Type 21

Turn movement data for surface street links are recorded in the record type 21. These data will change when detours are used.

4.2.2 FRESIM

(1) Record Type 29

A comprehensive freeway incident simulation procedure is provided in FRESIM. It is recommended by the user manual for work zone modeling. The user can specify either blockages or “rubbernecking” to occur on a lane-specify basis. The rubbernecking factor (in a percentage) represents the reduction in capacity for vehicles in remaining open lanes in the work zone area. Each incident occurs at the specified longitudinal position on a freeway link, extends over the user-specified length of the roadway, and last for any desired length of time (CORSIM Users’ Manual).

With the above function, it is convenient to set up: (a) number of closed lanes; (b) location of work zone (left, center or right of the road); (c) work zone length; (d) starting time of the work zone; (e) work zone duration; (f) location of the upstream warning sign for a work zone; (g) rubbernecking factor in the remaining open lanes in the work zone area.. FRESIM is therefore attractive for simulating work zone conditions due to this freeway incident specification function, which is defined in the record type 29.

Nevertheless, there are still some limitations in applications. For example, the time of the onset of the incident, which is measured from the start of the simulation, can not exceed 9999 seconds. The duration of an incident can not exceed 99999 seconds and the length affected by the incident cannot exceed 99999 feet.

(2) Record Type 20

Record Type 20 is used to record freeway link operation data. In this record type, the information contributing to work zone operation includes:

(a) Desired free-flow speed in a freeway link

This parameter specifies the desired, unimpeded, mean free-flow speed (in miles per hour) that is attained by traffic, in the absence of any impedance due to other vehicles, control devices or work zone activities. Under work zone conditions, the speed limit and driver’s compliance behavior in the freeway link may be changed. Then the free-flow speed may need to be varied.

The default value of this parameter is obtained from the simulator imported in Step 2. Users can reset this value considering the change of speed limit and driver’s compliance behavior in the work zone link.

(b) Car following sensitivity multiplier in a freeway link

The car following sensitivity multiplier permits users to adjust the car-following sensitivity on a link-by-link basis in a FRESIM network. The car following sensitivity factor represents a driver's desire to follow the preceding car. The value of car following sensitivity multiplier in a link contributes to the vehicle capacity in this link.

(3) Record Type 25

Turn movement data for surface street links are recorded in the record type 25. These data will be updated when detours are used.

4.3 Evaluation of Work Zone Plans in CORSIM

4.3.1 Evaluation Procedure

Microscopic traffic simulation models based on CORSIM is a powerful tool to evaluate pre-specified work zone operations based on detailed representations of traffic characteristics, network geometry characteristics, and traffic control plans.

To evaluate a work zone plan, there are five steps to follow:

Step 1: Build simulation model of the study roadway network in CORSIM

In the first step, a dataset describing the study roadway network should be defined for CORSIM. Geometrics of the network, traffic data such as volumes and turn movements, traffic control parameters such as sign and signal at intersections, and other information should be provided and can be recorded in a CORISM input file in format of TRF file.

Step 2: Specify work zone characteristics.

Major work zone characteristics include: (1) Lane closure alternative (number of closed lanes, number of usable counter flow lanes, number of access lanes); (2) Operation characteristics (work zone length, work starting time, work duration); (3) Work time and cost parameters used to evaluate work zone duration and maintenance cost; (4) Detour information (number of available detours and detour routes).

Step 3: Prepare simulation input file according to the work zone characteristics.

According to the work zone characteristics, the CORSIM input file defined in the first step is updated by changing existing records which contain work zone related information or adding new records which describe a lane closure activity. By this means, a new input file is generated. In this input file, work zone characteristics have been input into the study network.

Step 4: Run simulation.

After finishing creating an input file which interpreting work zone information into record types that can be recognized by CORISM, we can process the input file with the microscopic traffic simulator CORSIM.

The original input file without work zone information is also necessary to be processed with CORSIM because the net effect due to work zone should be obtained from the difference between simulation results with and without work zones. In order to reduce the statistical variance in simulation analysis, multiple simulation replications must be run with different random number seeds. The running time of one simulation iteration depends on scope of the size of the network, the number of time periods, the number of multiple runs.

Step 5: Evaluate Work Zone Delay

When CORISM terminates, an output file in format of OUT file will be generated. The CORSIM output file consists of cumulative NETSIM statistic data, specific NetSim Statistics data, cumulative FRESIM statistic data, FRESIM network statistics and network-wide Average Statistics for each time period. Various measures of effectiveness (MOE's), such as speeds, densities and delay time, can be calculated from the CORSIM output.

In our project, we use the net user delay due to work zones as our major MOE. Since we intend to evaluate the work zone effect from the system point of view, the "Delay Time" in Network-wide average statistics is used as the data to calculate the work zone delay (veh.hr). The work zone delay is estimated as:

$$ND_{\text{work zone}} = D_{\text{network delay with work zone}} - D_{\text{network delay without work zone}} \tag{4-1}$$

where $ND_{\text{work zone}}$ is the net user delay caused by work zone activities; $D_{\text{network delay with work zone}}$ represents the delay time in Network-wide average statistics for the last time period from the output of the simulation file with work zone information; $D_{\text{network delay without work zone}}$ represents the delay time in Network-wide average statistics for the last time period from the output of the original simulation file without work zone information.

The summary procedure for work zone plan evaluation in CORSIM is shown in Figure 4.1.

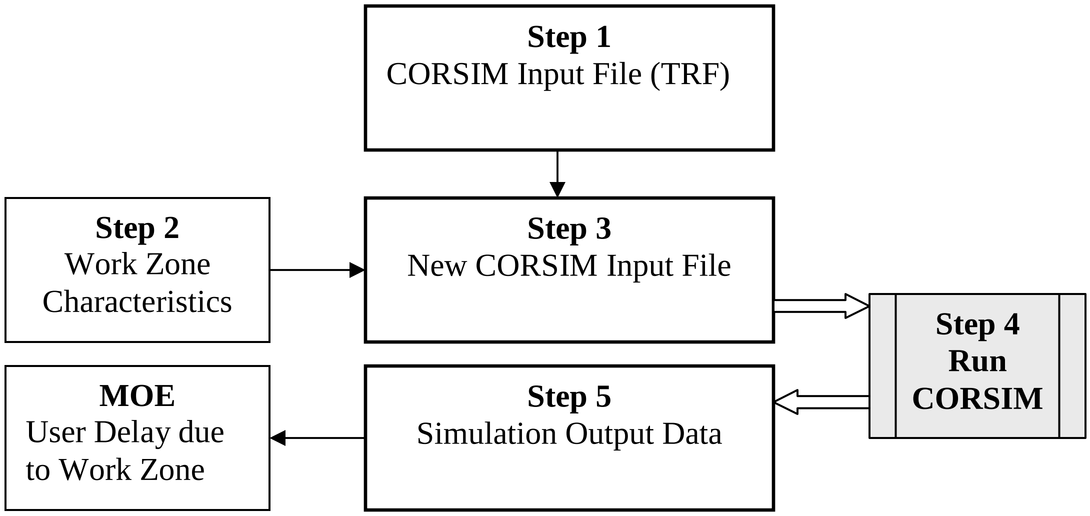


Figure 4.1 Procedure of Work Zone Plan Evaluation in CORSIM

4.3.2 Limitations in CORSIM

It must be noted that there are some limitations in CORSIM which may cause difficulties in work zone simulation with CORSIM.

1. Limitations Related to Input Data

(1) Simulation time

CORSIM can simulate up to 19 time periods with maximum duration, 9999 seconds, in each time period. Thus, the total simulation time cannot exceed 52.7 hours. Hence, we cannot simulate a work zone whose duration exceeds 52.7 hours in one CORSIM input file (TRF file).

Due to the first limitation, several TRF files instead of one TRF file should be needed to simulate a work zone with long-term duration. The delay time is the sum of the results from the corresponding output files.

(2) The onset time of an incident in record type 29

For record type 29, which is used to simulate freeway work zones, CORSIM only allows users to specify the onset time of an incident (in seconds) at up to 9999 seconds, (Time is measured from the start of the simulation in CORSIM.) This indicates that the start time of the simulation has to occur less than 9999 seconds ahead of work zone starting time and two zones cannot be successive in a TRF file if the first zone's duration exceeds 9999 seconds.

Due to the second limitation, the simulation start time should vary according to the work zone starting time. For example, if the work zone starting time is 9:00 am, the earliest simulation start time will be 6:13 am.

2. Limitations Related to Output Data

(1) Vehicles entering the study network

We noted that CORSIM has difficulty dealing with storage of vehicles on short, congested links. In the congested network, once the queues extend back to the entrance node and block vehicles from entering the network at their scheduled time. Vehicles that were scheduled to depart were not able to do so. The "departure delays" of those vehicles backed up behind entrance nodes will not be included in the total delay estimates in output statistics.

This limitation may result in underestimating of user delays in over-saturated conditions. That is to say, CORSIM may be unable to provide precise delay estimations for those "bad" work zone plans that may cause queue spillback.

(2) Vehicles leaving the study network

Since FRESIM only reports delay for the vehicles that have arrived at their destinations during the analysis period, a potential underestimation of travel time may happen if a queue has not be cleared or there are vehicles still on their way when the simulation ends. To solve this problem, additional simulation periods may have to be added to clear the queue. This also indicates that, if several TRF files are needed to simulate a work zone with long-term duration, the end time of simulation of each TRF file should be set to avoid loosing the information about vehicles stuck in a queue, such as in an off-peak hour.

4.4 Comparison of Microscopic Simulation and Analytic Methods

In a macroscopic traffic flow model, speed is derived from the relation among flow, speed and density, while in a microscopic simulation model speed is derived from the car following theory. In this subsection, we discuss the difference and the relation between microscopic car-following models in CORSIM and macroscopic traffic flow models used in our analytic methods.

The macroscopic traffic flow models identify the relation between the three traffic flow parameters, namely flow (*Q*), speed (*V*), and density (*K*), which can be measured fairly easily in the field using standard loop detectors or traffic counters.

Unlike macroscopic models describing the behavior of a stream of vehicles along a roadway stretch, microscopic car following models describe the behavior of a pair of vehicles within a traffic stream. Steady-state microscopic car-following models characterize the relation between the vehicle's desired speed and the distance headway between the lead and follower vehicles.

4.4.1 Car-Following Model in CORSIM

In FRESIM and NETSIM, the Pipes car-following model serves as the basis for steady-state car-following logic (Crowther, 2001). In Pipes' model, the distance headway has a linear relation with speed. The car-following behavior of a vehicle is constrained by a maximum speed, which is commonly known as the free flow speed.

(1) FRESIM

The FRESIM model utilizes the Pitt car-following behavior that was developed by the University of Pittsburgh (Halati et al., 1997).

$$\begin{aligned} H &= H_j + c_3 V + b c_3 \Delta V^2 \\ &= H_j + c_3 V \qquad \text{where, } \Delta V = 0 \text{ under steady-state conditions} \end{aligned} \tag{4-2}$$

$$V = \min\left(\frac{H - H_j}{c_3}, V_f\right) \tag{4-3}$$

where H = distance headway between lead and follower vehicles (km)

H_j = jam density headway (km)

c_3 = driver sensitivity factor (hr)

V = follower vehicle speed (km/hr)

b = calibration constant which equals 0.1 if the speed of the follower vehicle exceeds the speed of the lead vehicle, otherwise it is set to 0.

$\triangle V$ = difference in speed between lead and follower vehicle (km/hr)

V_f = free-flow speed (km/hr)

The FRESIM model utilizes 10 driver types, which are characterized by driver sensitivity factors ranging from 0.6 to 1.5.

(2) NETSIM

The car-following model in NETSIM incorporates a driver reaction time and the ability of vehicles to decelerate at feasible rates without resulting in vehicle collisions.

$$H = H_j + \Delta S + \Delta R + S_F - S_L$$

$$= H_j + V\Delta t \quad \text{under steady-state conditions}$$

$$= H_j + V\frac{1}{3600} \quad \text{where } \Delta t = 1 \text{ seconds in NETSIM} \tag{4-4}$$

$$V = \min\left(\frac{H - H_j}{c_3}, V_f\right) \tag{4-5}$$

where,

H = distance headway between lead and follower vehicles (km)

H_j = jam density headway (km)

$\triangle S$ = distance traveled by follower vehicle over time interval $\triangle$t (km)

$\triangle R$ = distance traveled by follower vehicle during its reaction time (km)

S_F = distance traveled by follower vehicle to come to a complete stop (km)

S_L = distance traveled by lead vehicle to come to a complete stop (km)

V_f = free-flow speed (km/hr)

4.4.2 Relation between Microscopic and Macroscopic Models

The car-following model of traffic has a harmonious tie-in to macroscopic theory. The following procedures can integrate the two approaches.

In traffic flow theory, the relation among flow (Q), density (K), and speed (V) is:

$$Q = KV \tag{4-6}$$

Assuming all vehicles in the traffic stream maintain the same headway distance, we obtain:

$$H = \frac{1}{K} \tag{4-7}$$

For Pipes' car following model, we substitute Eq.(4-7) into Eq.(4-5) and Eq. (4-6). Then this microscopic car-following model can be related mathematically to the macroscopic speed-density relationship through the following forms:

$$V = \min\left(\frac{H - H_j}{c_3}, V_f\right) = \min\left(\frac{\frac{1}{K} - \frac{1}{K_j}}{c_3}, V_f\right) \tag{4-8}$$

$$Q = KV = \frac{1}{H}V = \frac{V}{H_j + c_3 V} = \frac{V}{\frac{1}{K_j} + c_3 V} \tag{4-9}$$

The traffic stream model that evolves from the Pipes car-following model is **multi-regime** in the sense that a different model is utilized for the congested versus uncongested regimes. Specifically, the Pipes model assumes that the traffic stream speed is insensitive to the traffic density in the uncongested regime.

Greenshield's model developed in 1934 is the most famous macroscopic traffic stream model. It is applied in the analytic work zone delay estimation method presented in Chapter 3. In this model, speed is a linear function of density.

$$V = V_f - \frac{V_f}{K_j}K \tag{4-10}$$

where V_f = free-flow speed;

K_j = jam density.

Substituting Eq.(4-10) into Eq.(4-6), we obtain:

$$Q = KV = \frac{K_j}{V_f}(V_f - V)V = K_j V - \frac{K_j}{V_f}V^2 \tag{4-11}$$

From Eq.(4-11), we see that Greenshields' traffic flow model is a **single-regime** model.

To show the difference of Pipes' model and Greenshield's model, we give two numerical examples in this subsection. The relations among flow (Q), density (K), and speed (V) in the two models are illustrated in Figure 4.2, given K_j =80, V_f =80, c_3=1/3600. Another numerical example (Figure 4.3) shows the same comparison when the capacity (Q_c) is the same in both models.

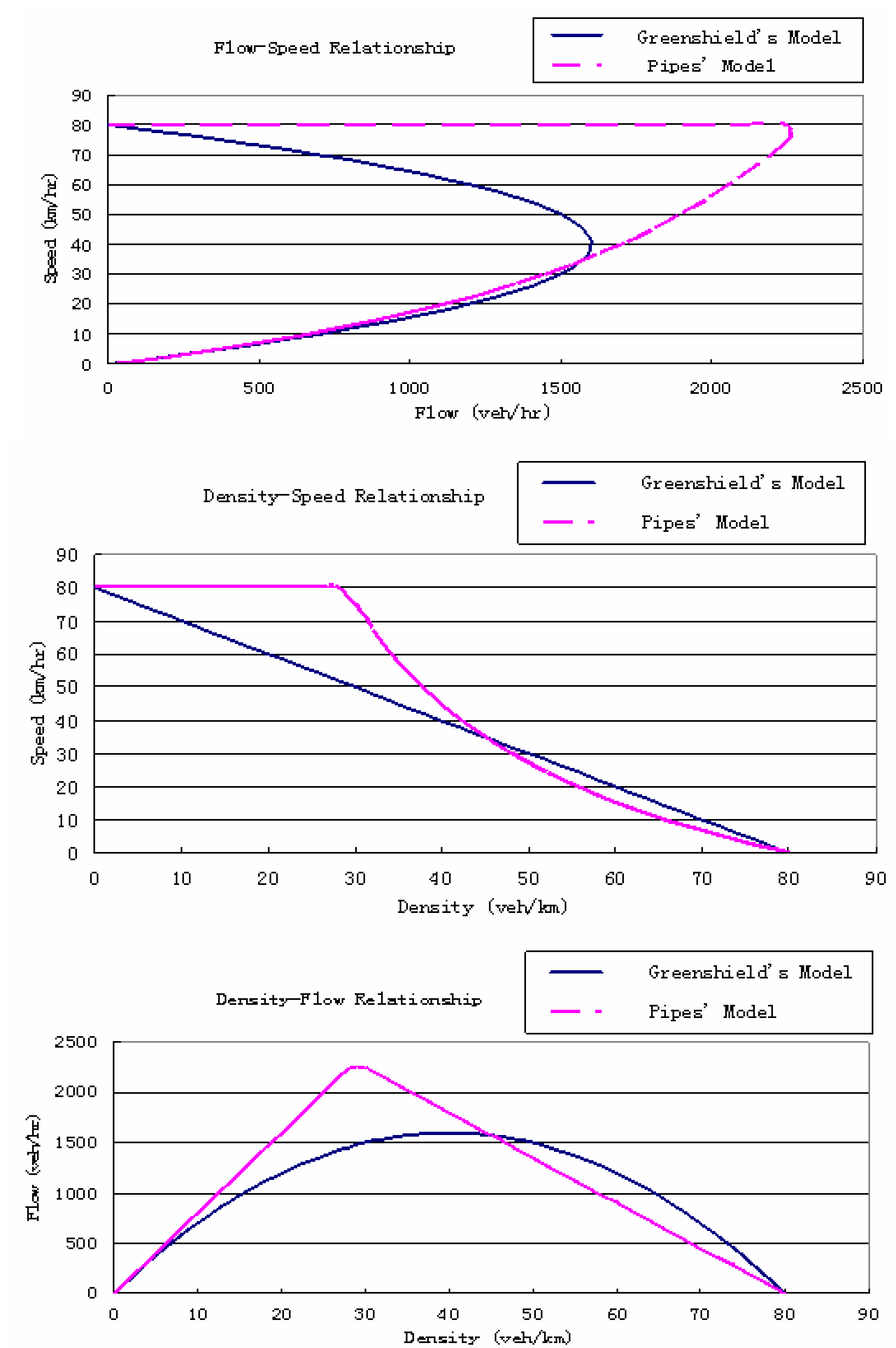


Figure 4.2 Comparison of the Two Models (K_j=80, V_f=80, c_3=1/3600)

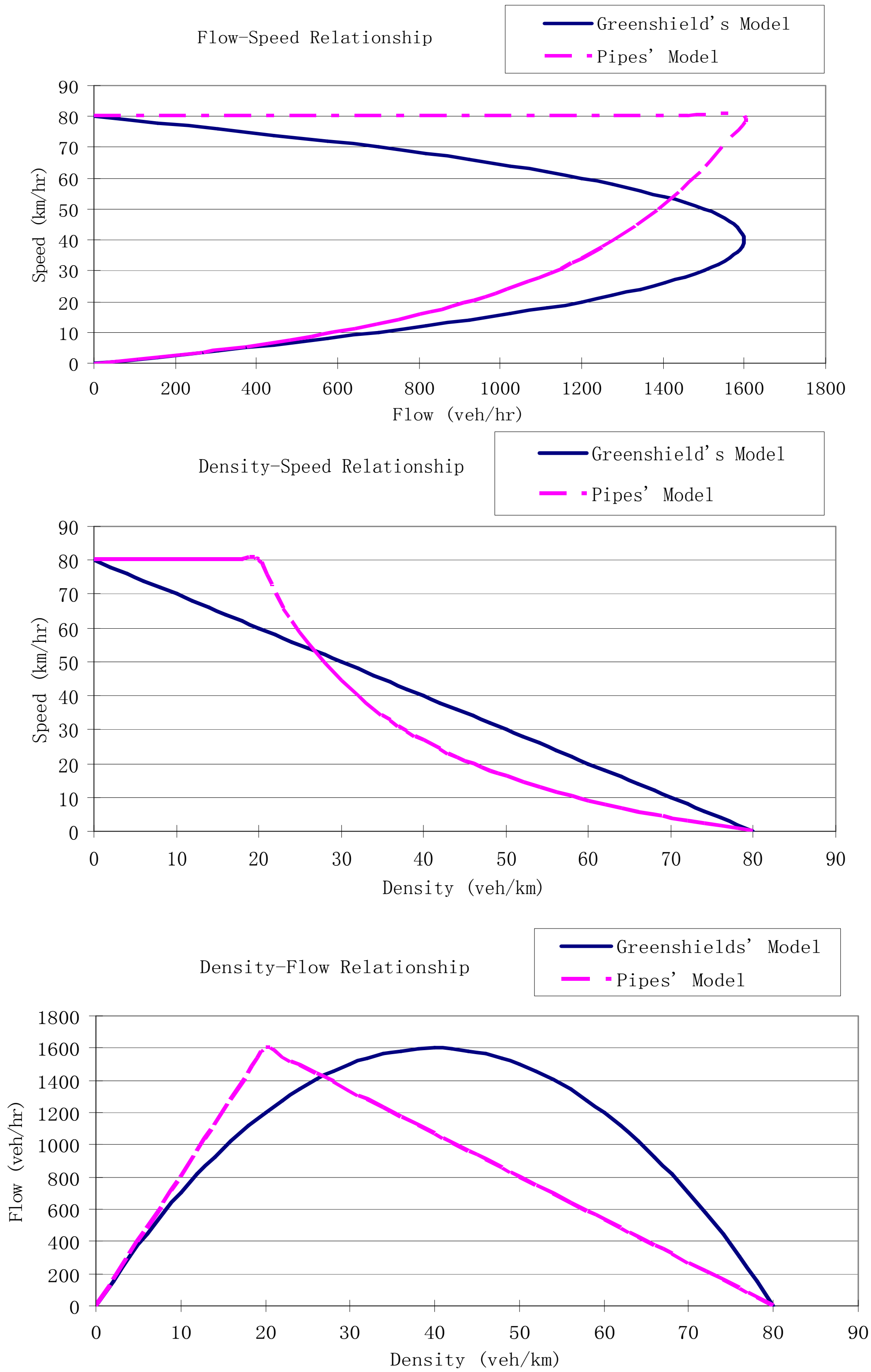


Figure 4.3 Comparison of the Two Models (K_j=80, V_f=80, Q_c=1800)

4.5 Numerical Example

In this numerical example, we use the CORSIM simulation model to estimate user delays caused by a work zone in a two-lane one-way freeway segment in different traffic conditions. The results are compared to user delay calculated with analytic method.

4.5.1 Test Network Configuration

A hypothetic work zone site is modeled in FRESIM in a 3-mile freeway segment (Figure 4.4). The freeway segment consists of three links, 1-mile upstream work zone link, 1-mile reduced-lane section and 1-mile downstream work zone link. A work zone with one-lane closure will be in the middle link.

Since the simulation model is hypothetical, there are no field data available to calibrate it. Here, we modified two parameters, the car following sensitivity factor and rubbernecking factor, to specify the roadway capacities outside and inside the work zone segment given as inputs. Here the capacity is defined as the maximum hourly flow passing through the work zone, which can be obtained by gradually increasing inflow rate until the outflow rate keeps unchanged.

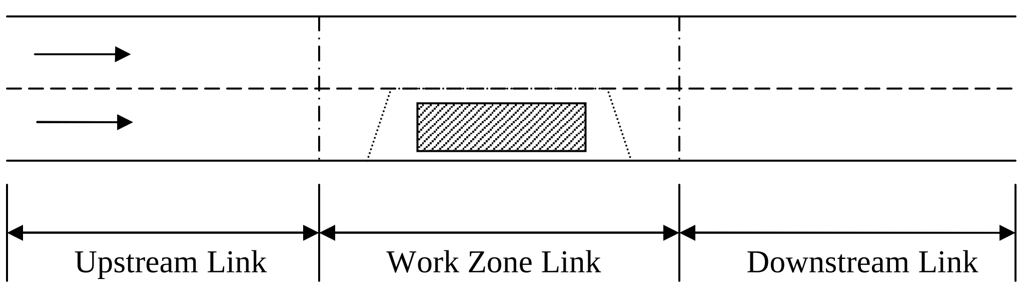


Figure 4.4 Test Network Configuration

Keeping default value of the car following sensitivity factor, we get the roadway capacity of 2200 vph. The rubbernecking factor in the remaining lane is set as 45% to result in a work zone capacity of 1300 vph. The free flow speed is set to be 65 mph in all three links. All other parameters such as driver characteristics use default values provided by FRESIM.

To compare simulation results with analytic solutions, the following values are specified: average work zone speed below capacity of 55 mph, speed at capacity of 45 mph, jam density 200 veh/mile.

4.5.2 Sensitivity Analysis

(1) Steady Traffic Inflows

Assume that the work zone duration is 1 hour, the work zone length is 0.5 mile, the location of upstream warning sign is 5000 ft, and the distance from upstream node is 0 ft. The simulation time is set to 4 hours and the work zone starts 1 hour after the start of simulation. In this example, the traffic inflows are steady for the first 3 hours. The traffic inflow in the last hour is set to be zero to clear the vehicles in the network.

User Delay vs. V/C Ratio

Scenarios with increasing traffic inflows are tested. Each work zone scenario is run 10 times for each volume to work zone capacity ratio ranging from 0.5 to 2.5. Table 4.1 and Figure 4.5 show the averages net user delay obtained from CORSIM and the net delay calculated by deterministic analytic method for all scenarios.

Table 4.1 User Delay for Traffic Inflows for V/C Ratios from 0.5 to 2.5

V/C	Volume (vhp)	Capacity (vph)	Avg. Delay (veh.hr) (CORSIM)	Standard Deviation	Net Delay (veh.hr) (Analytic Method)
0.5	650	1300	0.74	0.05	0.91
0.6	780	1300	1.06	0.06	1.09
0.7	910	1300	1.46	0.08	1.27
0.8	1040	1300	2.26	0.12	0.67
0.9	1170	1300	5.63	0.55	0.75
1.0	1300	1300	48.17	3.14	4.45
1.1	1430	1300	119.60	4.68	71.37
1.2	1560	1300	197.37	3.10	142.13
1.3	1690	1300	288.87	5.27	216.73
1.4	1820	1300	385.42	6.07	295.17
1.5	1950	1300	490.53	4.32	377.46
1.6	2080	1300	583.67	7.10	463.58
1.7	2600	1300	655.21	6.48	553.50
1.8	2340	1300	712.16	13.33	647.00
1.9	2470	1300	756.14	14.62	745.00
2.0	2600	1300	800.59	22.67	846.49
2.1	2730	1300	844.02	21.01	951.50
2.2	2860	1300	875.47	12.60	1060.99
2.3	2990	1300	895.14	20.21	1174.00
2.4	3120	1300	906.95	19.26	1293.86
2.5	3250	1300	922.05	29.95	1411.55

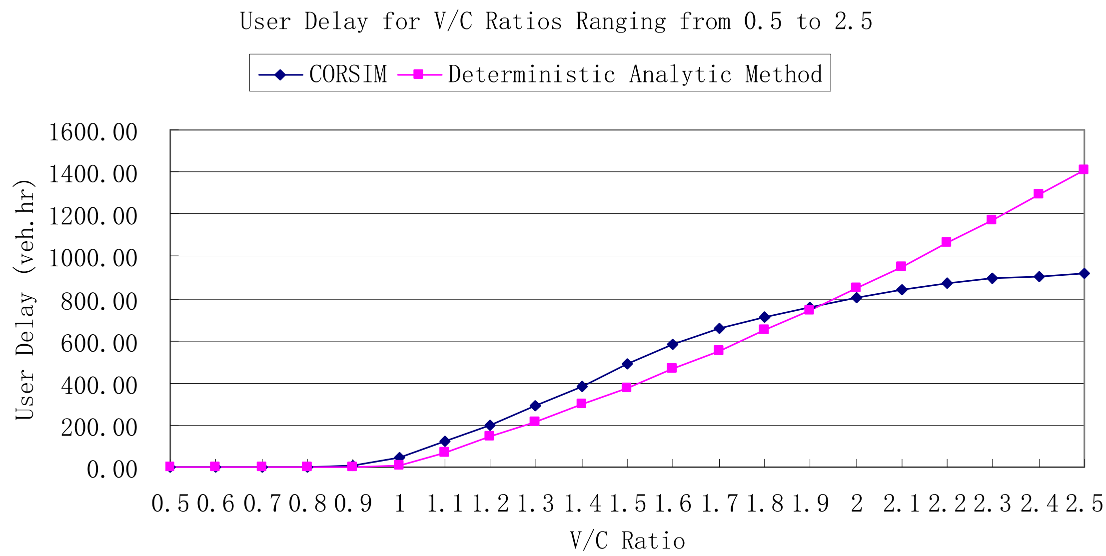


Figure 4.5 User Delay for V/C Ratios Ranging from 0.5 to 2.5

As seen from Table 3.1 and Figure 3.4, CORSIM estimates higher user delays than the deterministic analytical results in most scenarios, which is expected.

However, at high Volume/Capacity ratios the CORSIM estimates become lower than analytical results and the run-to-run standard deviation becomes larger. After checking the output files and the corresponding animation file generated by TSIS, we find that the queue spills back from the upstream work zone link to the entrance node when the V/C ratio exceeds 2.0. As we have discussed in the previous subsection about CORSIM limitations, CORSIM ignores the vehicles which cannot enter the network due to queue spillback, thus underestimates of total delay time in over-saturated conditions.

(2) Time-Dependent Traffic Inflows

In this case, traffic inflows are time-varying in a day. Two scenarios with different traffic volumes are analyzed to test the simulation results in different traffic conditions. In the first scenario for uncongested network, the baseline hourly distribution of the traffic inflow, shown in Table 4.2, uses detector data obtained from station B2500 along northbound US 1 in Maryland. In the second scenario for congested network, hourly traffic volumes are 1.5 times higher than the baseline volumes.

The fixed work zone setup time and the average maintenance time are assumed to be 1 hour and 10 hour per lane mile, respectively. Then the work zone length can be derived from the work zone duration.

User Delay vs. Work Zone Start Time

For a 0.5-mile long work zone with one lane closure, the work duration needed to finish the work is 6 hours including 1 hour setup time and 5 hour maintenance time. Given the location of upstream warning sign as 5000 ft and the distance from upstream node to work zone start point as 5000 ft, we seek to perform a sensitivity analysis on the impact of work zone start time on user delay per lane-mile. User delays with different work zone start time in both scenarios are shown in Table 4.3, in which both simulation results and results from the analytic method presented in Chapter 3 are provided.

The results for Scenario 1 with baseline traffic volumes are shown in Figure 4.6. We can see that user delays increase sharply when the work zone start time approaches peak times, which results in lane closure in peak hours. Comparing simulation results and analytic results, we find the trend lines are almost the same. In low-volume hours, the analytic method estimates a little bit higher delays than does CORSIM due to probably overestimation of average work zone speed. In other times, the analytic method estimates lower delays than does CORSIM. Since the traffic condition is under-saturated in scenriao 1, the comparison result is consistent with that from the steady traffic inflow example, as we expected.

Figure 4.7 displays the results for Scenario 2 with higher traffic volumes. The changes of user delays become more sensitive to the work zone start time. The trend lines of simulation results and analytic results are still similar to each other. However, when work zone takes place in peak hours, delays obtained from CORSIM become fairly insensitive to work zone start times and they are much lower than delays calculated with the analytic method. As we have discussed in the steady traffic inflow case, this can be explained by the CORSIM'S inability to track the vehicles which can not enter the network due to queue spill back to entry nodes. Once the traffic condition becomes so over-saturated that the entry nodes are blocked, CORSIM may underestimate delay times because it does not consider the "departure delays" of those vehicles stuck outside the network.

Table 4.2 Hourly Traffic Distribution (US-1, North Bound, Station B2500)

Time Period	Time	Baseline Volume (vhp)	High Volume (vhp)
1	0:00-1:00	155	233
2	1:00-2:00	74	111
3	2:00-3:00	74	111
4	3:00-4:00	73	110
5	4:00-5:00	156	234
6	5:00-6:00	362	543
7	6:00-7:00	757	1136
8	7:00-8:00	1225	1838
9	8:00-9:00	1270	1905
10	9:00-10:00	962	1443
11	10:00-11:00	960	1440
12	11:00-12:00	1016	1524
13	12:00-13:00	1271	1907
14	13:00-14:00	1248	1872
15	14:00-15:00	1247	1871
16	15:00-16:00	1451	2177
17	16:00-17:00	1623	2435
18	17:00-18:00	1662	2493
19	18:00-19:00	1318	1977
20	19:00-20:00	867	1301
21	20:00-21:00	633	950
22	21:00-22:00	521	782
23	22:00-23:00	372	558
24	23:00-24:00	231	347
AADT		19528	29292
Average Hourly Volume		813	233

Table 4.3 User Delay for Baseline and High Traffic Volumes with Different Work Zone Start Times

Work Zone Start Time	User Delay (Baseline Volume) (veh.hr/lane.mile)		User Delay (High Volume) (veh.hr/lane.mile)	
	CORSIM	Analytic Result	CORSIM	Analytic Result
0	1.42	6.11	1.48	9.16
1	2.84	10.22	9.92	15.34
2	2.84	18.1	702.42	1097.47
3	140.76	26.28	2469.76	3389.6
4	187.82	32.36	4258.44	5967.74
5	191.24	37.85	6021.46	8825.03
6	191.24	42.32	6021.46	12128.21
7	223.24	45.84	9795.44	16639.33
8	328.7	46	9858.74	14775.33
9	479.76	45.84	8166.36	12087.33
10	965.16	350.15	8446.84	18329.33
11	2488.14	1300.48	9279.56	26719.33
12	4230.22	2974.42	10184.8	34839.33
13	5280.18	4682.62	10138.68	36399.33
14	6790.04	5524.97	10065.86	36589.33
15	7542.84	5520.77	10204.6	34809.33
16	5751.64	4005.45	10287.2	28565.33
17	2409.66	1518.14	10182.56	15473.33
18	193.84	62.831	5047.54	3398.188
19	5.66	19	178.42	28.48
20	2.72	13.58	7.16	20.36
21	1.44	9.76	3.74	14.63
22	0.68	6.69	1.64	10.03
23	0.3	5.22	0.78	7.82

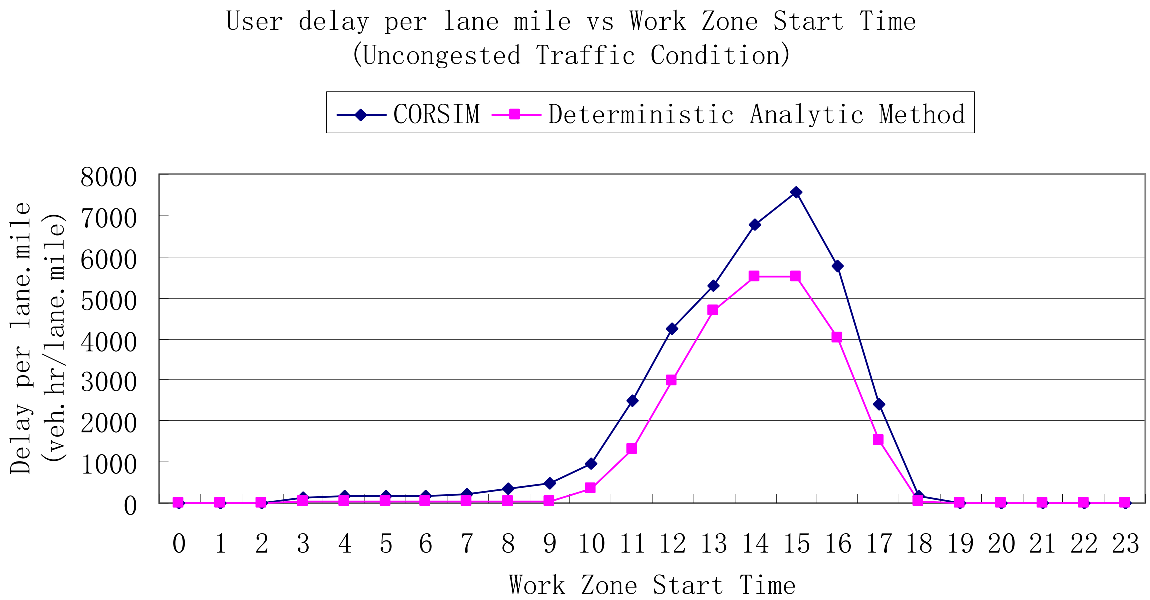


Figure 4.6 User Delays with Different Work Zone Start Time
(Scenario 1, Baseline Volumes, Work Duration=6 hours)

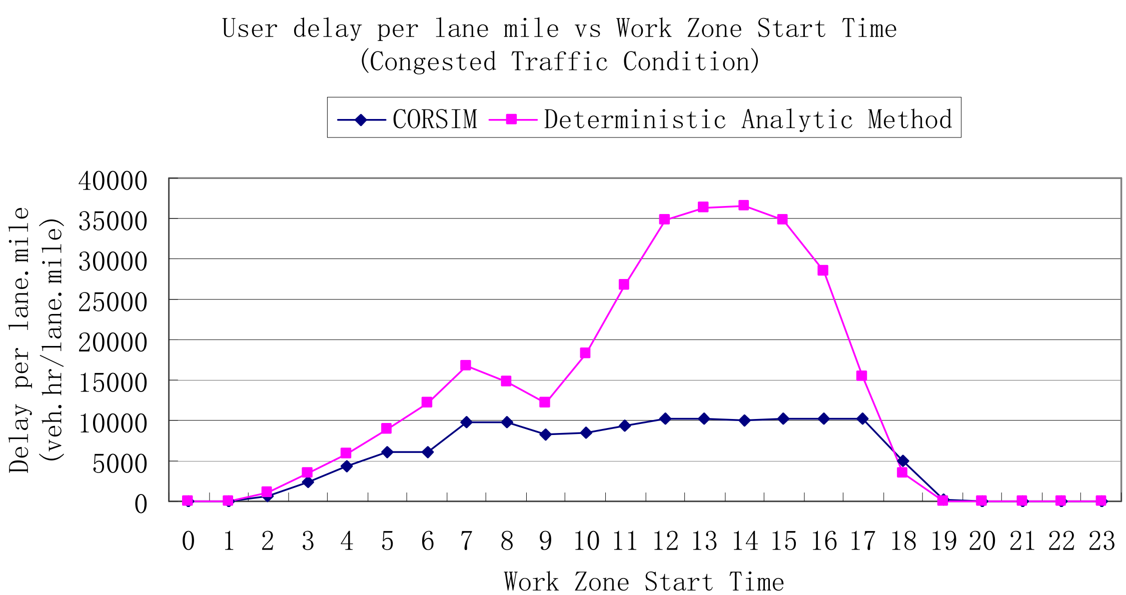


Figure 4.7 User Delays with Different Work Zone Start Time
(Scenario 2, High Volumes, Work Duration=6 hours)

User Delay vs. Work Zone Duration

Assuming that the work zone start time is 23:00 pm, we test the impact to work zone duration, which is related to work zone length, on the user delays caused by per lane mile work zone activity. The results are displayed in Table 4.4.

Figure 4.8 and Figure 4.9 show the user delays with different work zone duration with baseline traffic volumes and with high traffic volumes, separately. The trend lines of the change of user delays from CORSIM and from the analytic method are consistent with each other. Still, CORSIM results are higher than analytic results in Scenario 1 with base line volumes under uncongested traffic conditions, and are lower than analytic results in Scenario 2 with high volumes under congested traffic conditions, especially when closing lanes in peak hours.

Table 4.4 User Delays for Baseline and High Traffic Volumes with Different Work Zone Durations

Work Zone Duration	Work Zone Length (mile)	User Delay(Baseline Volume) (veh.hr/lane.mile)		User Delay (High Volume) (veh.hr/lane.mile)	
		CORSIM	Analytic Result	CORSIM	Analytic Result
2	0.1	0.9	2.64	1.9	3.95
3	0.2	0.45	3.14	1	4.71
4	0.3	0.33	3.65	0.77	5.47
5	0.4	0.2	4.15	0.55	6.22
6	0.5	0.3	5.21	0.78	7.81
7	0.6	1.13	7.69	1.25	11.53
8	0.7	1.9	12.86	6.99	19.29
9	0.8	19.78	21.24	416.4	699.43
10	0.9	63.67	29.93	1370.8	1902.622
11	1	63.9	36.51	2129.51	3009.956

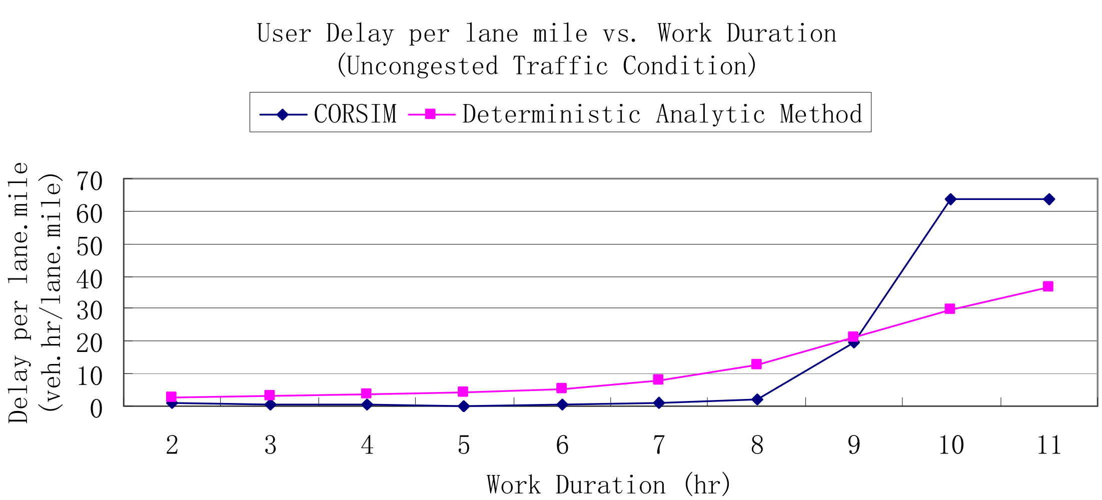


Figure 4.8 User Delays with Different Work Zone Duration
(Scenario 1, Baseline Volumes, Work Zone Start Time=23:00)

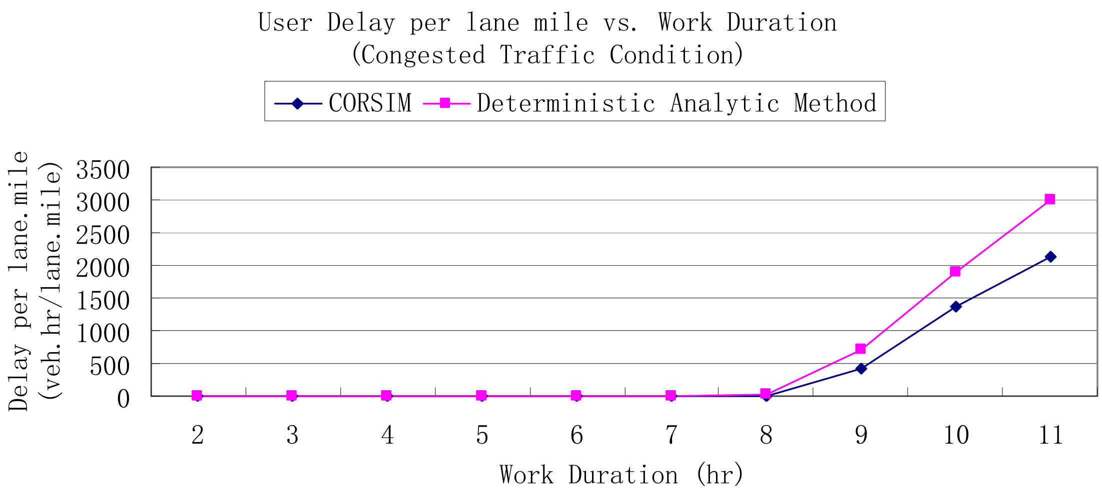


Figure 4.9 User Delays with Different Work Zone Duration
(Scenario 2, High Volumes, Work Zone Start Time=23:00)

4.6 Summary

In this chapter, work zone plans are analyzed with simulation and user delays are estimated using CORSIM software. The incident function in FRESIM, a freeway component in COSRIM, is chosen to simulate the work zone sites. The comparison of microscopic car-following model and macroscopic traffic flow model is discussed. Sensitivities of the delays estimated by CORSIM and by the analytic method presented in Chapter 3 with respect to traffic conditions and work zone characteristics are analyzed. It is found that work zone activity may significantly affect traffic conditions and cause high user delays if the work zone characteristics, such as work zone start time, work zone duration, are not properly planned. It is also found that CORSIM estimates higher delays than the analytic method under uncongested traffic conditions while estimating lower delays than the analytic method does under congested traffic conditions. This can be explained by the inability of CORSIM to calculate the delays of the vehicles that cannot enter the network as scheduled due to the queue spillbacks to traffic entry nodes in an over-saturated road network.

Chapter 5 Work Zone Optimization based on Analytic Model for Steady Traffic Inflows

In this chapter, work zone optimization models for steady traffic inflows are developed for two-lane highway and multiple-lane highway work zones. The optimization problem statement is discussed in Section 5.1. Section 5.2 presents an optimization model formulation without considering a time-cost tradeoff while Sections 5.3 and 5.4 formulate an optimization model considering a time-cost tradeoff. A numerical example is provided in Section 5.5.

5.1 Problem Statement

Optimization can be defined as the process of finding the conditions that give the maximum or minimum value of a function. To describe an optimization problem, we should define: (1) the objective function; (2) the decision variables; and (3) inequality or equality constraints.

In Chapter 3, we discussed important work zone characteristics affecting work zone costs. These characteristics includes: (1) lane closure alternatives; (2) operation characteristics such as work zone length and work zone schedule; (3) time and cost parameters; and (4) detour type.

In this work zone optimization problem, the objective function is to minimize the total cost per maintained lane length. The cost functions in terms of the above characteristics have been obtained in Chapter 3.

In this chapter, we present two optimization models. In one model, the work zone length is used as the decision variables and the objective function is shown in Eq.5-1. In another optimization model, the work zone length and work rate, which represents tradeoff between work time and cost, are optimized simultaneously. Its objective function is shown in Eq.5-2.

$$\text{Min } Ct = f\,(\text{work zone length}) \tag{5-1}$$

$$\text{Min } Ct = f\,(\text{work zone length, work zone rate}) \tag{5-2}$$

For steady traffic inflows, closed-form objective function can be obtained. The classical methods of differential calculus can be used to find the unconstrained minima of the function of certain decision variables.

5.2 Work Zone Alternatives

5.2.1 Review of Work Zone Characteristics

The following variables are defined for describing work zone characteristics:

n_1 = the number of lanes in the Direction 1

n_2 = the number of lanes in the Direction 2

n_{c1} = the number of closed lanes in the Direction 1

n_{c2} = the number of closed lanes in the Direction 2

For two-lane two-way highways, n_{c2} is a binary variable. If the lane in the opposite direction is open for traffic in both directions, n_{c2}=0. If the opposite lane is also closed, n_{c2}=1 and in such cases the traffic flows in both directions have to be fully diverted to detour routes.

For multiple-lane two-way highways, n_{c2} represents the number of usable counter-flow lanes for the traffic along the original direction when a crossover is used.

n_{a1} = the number of lanes used to access among the closed lanes

n_{o1} = the number of open lanes open for the traffic inflow in Direction 1

n_{o2} = the number of open lanes open for the traffic inflow in Direction 2

n_w = the number of maintained lanes

For two-lane two-way highways,

$$n_{o1} = n_1 - n_{c1} + (1 - n_{c2}) \quad (5\text{-}3a)$$

$$n_{o2} = n_2 - n_{c2} \quad (5\text{-}3b)$$

$$n_w = n_{c1} + n_{c2} - n_{a1} \quad (5\text{-}3c)$$

For multiple-lane two-way highways,

$$n_{o1} = n_1 - n_{c1} + n_{c2} \quad (5\text{-}4a)$$

$$n_{o2} = n_2 - n_{c2} \quad (5\text{-}4b)$$

$$n_w = n_{c1} - n_{a1} \quad (5\text{-}4c)$$

L_w = the work zone length;

D_w = the work zone duration;

z_1 = the fixed setup cost per zone;

z_2 = the average additional cost per lane length;

z_3 = the fixed setup time per zone;

z_4 = the average additional time per lane length;

C_w = work zone capacity;

C_0 = normal capacity;

V_w = average work zone speed;

Q_1 = traffic inflow in Direction 1 (veh/hr)

Q_2 = traffic inflow in Direction 2 (veh/hr)

p_1 = the diverted fraction for the traffic inflow Q_1;

p_2 = the diverted fraction for the traffic flow Q_2;

Q_{1m} = the traffic flow in direction 1 along work zone link;

$$Q_{1m}= (1-p_1)\, Q_1 \tag{5-5}$$

Q_{2m} = the traffic flow in direction 2 along counter work zone link;

$$Q_{2m}= (1-p_2)\, Q_2 \tag{5-6}$$

L_{AB} = the length of mainline segment AB, shown in Figure 5.1;

L_{AC} = the length of detour segment AC, shown in Figure 5.1;

L_{CD} = the length of detour segment CD, shown in Figure 5.1;

L_{DB} = the length of detour segment DB, shown in Figure 5.1;

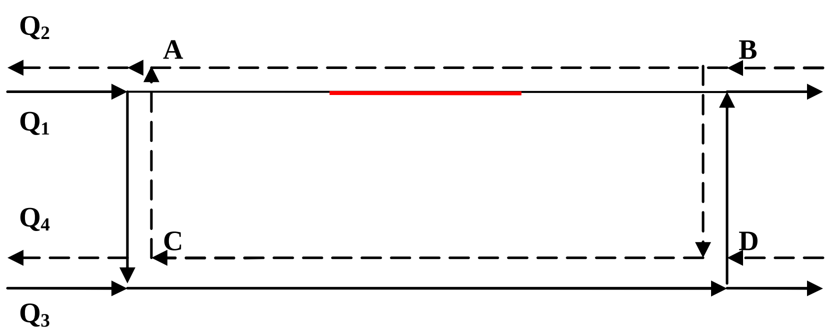


Figure 5.1 Geometric of the Study Network for Steady-Flow Traffic Inflows

In this optimization problem, work zone length and time-cost parameters are chosen to be decision variables. However, other work zone characteristics, such as lane closure alternatives, are also significant elements. Several alternatives are specified for two-lane two-way highway work zones and four-lane two-way highway work zones, which are typical examples of multiple-lane two-way highway work zones. Optimization models are developed for each alternative.

5.2.2 Two-Lane Two-Way Highway Work Zone Alternatives

(1) Alternative 2.1: One lane in Direction 1 is closed. The traffic inflows from both directions are alternated on another open lane. No detour is available. In this case, n_{c1}=1, n_{c2}=0,

$p_1=0\%$ and $p_2=0\%$.

(2) Alternative 2.2: A fraction of traffic flow in Direction 1 is diverted to a detour. Traffic flow in Direction 2 is not diverted. In this case, n_{c1}=1, n_{c2}=0, 100%>$p_1>0\%$ and $p_2=0\%$.

(3) Alternative 2.3: All traffic flow in Direction 1 is diverted to the detour while the remaining lane is only used for traffic in the other direction. In this case, n_{c1}=1, n_{c2}=0, $p_1=100\%$ and $p_2=0\%$.

(4) Alternative 2.4: All traffic in both directions is diverted to the alternate route and both lanes are closed for work. In this case, n_{c1}=1, n_{c2}=1, $p_1=0\%$ and $p_2=0\%$

The geometries of all four alternatives are shown in Figure 5.2.

5.2.3 Four-Lane Two-Way Highway Work Zone Alternatives

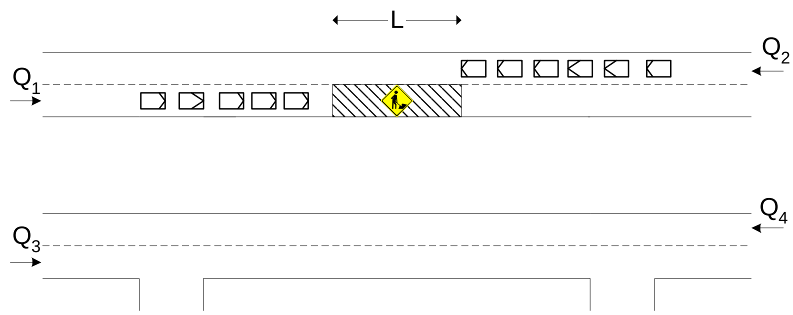
(1) Alternative 4.1: One of the two lanes in Direction 1 is closed for Q_1 traffic. No fraction of Q_1 traffic is detoured. Q_2 traffic is not impacted. In this case, n_{c1}=1, n_{c2}=0 and $p_1=0\%$.

(2) Alternative 4.2: One of the two lanes in Direction 1 is closed for Q_1 traffic. A fraction of Q_1 is detoured. Q_2 traffic is not impacted. In this case, n_{c1}=1, n_{c2}=0 and $p_1>0\%$.

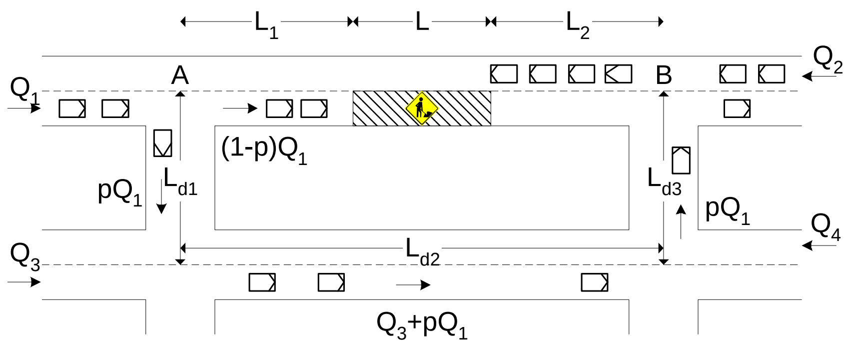
(3) Alternative 4.3: Both of the two lanes in Direction 1 are closed. All Q_1 traffic is detoured. Q_2 traffic is not impacted. In this case, n_{c1}=2, n_{c2}=0 and $p_1=100\%$.

(4) Alternative 4.4: Both of the two lanes in Direction 1 are closed. All Q_1 traffic crosses over into one lane in Direction 2. Q_2 traffic is impacted. In this case, n_{c1}=2, n_{c2}=1 and $p_1=0\%$.

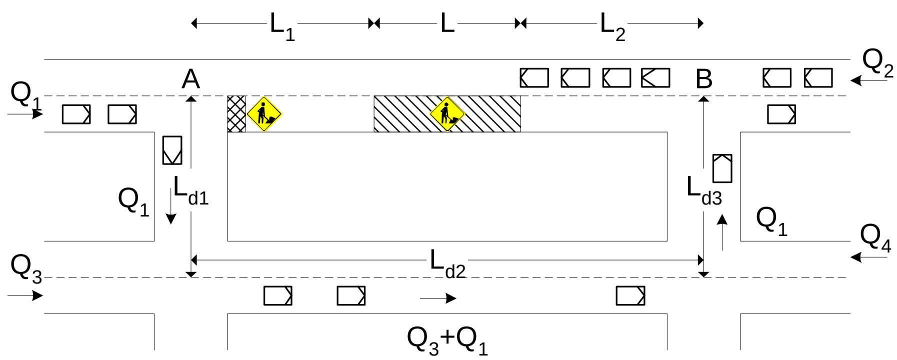
The geometries of all four alternatives are shown in Figure 5.3.



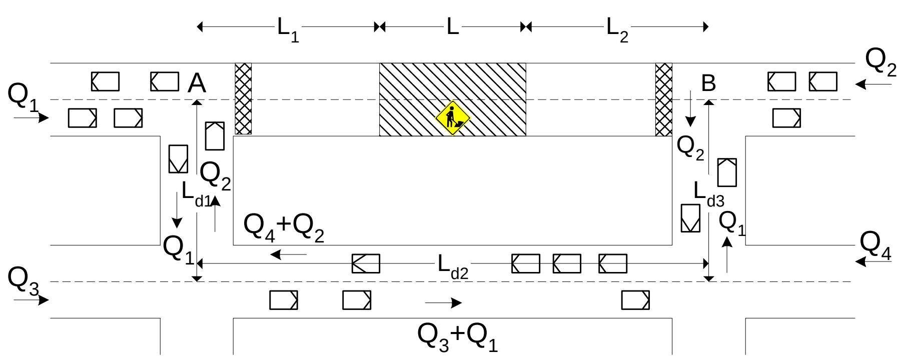
(a) Alternative 2.1: without Detour



(b) Alternative 2.2: with Fraction through Detour

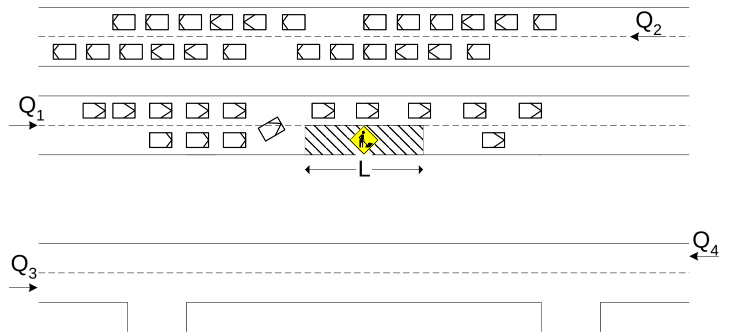


(c) Alternative 2.3: Detour for Only One Direction

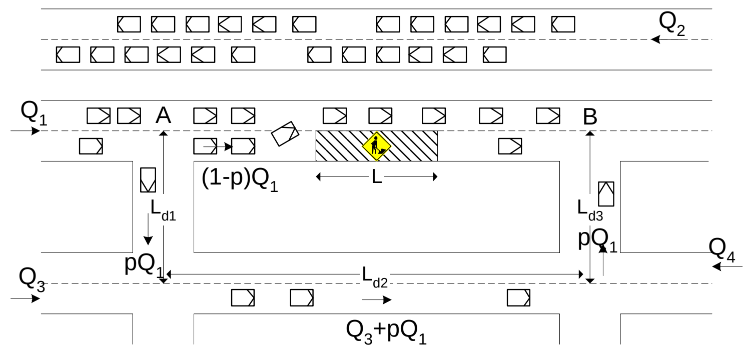


(d) Alternative 2.4: Two Directions Detoured

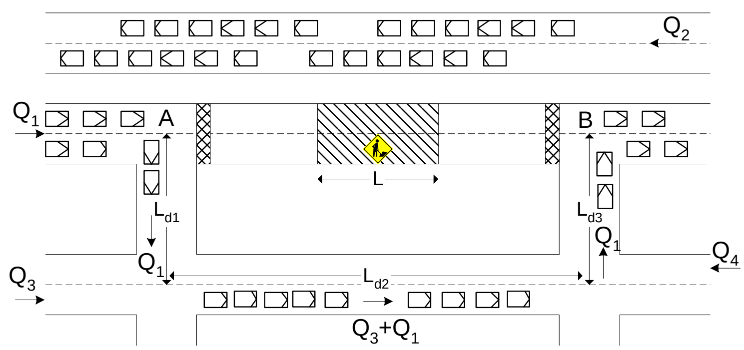
Figure 5.2 Geometries of Analyzed Work Zones for Two-Lane Two-Way Highways



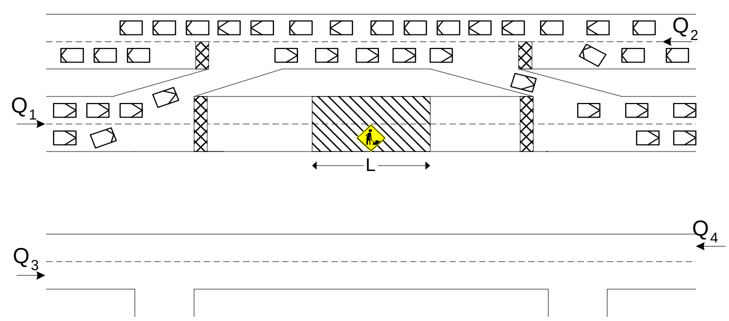
(a) Alternative 4.1: No Detour, One of the Two Lanes closed for Q_1 Traffic



(b) Alternative 4.2: A Fraction of Q_1 Traffic through Detour



(c) Alternative 4.3: All Q_1 through Detour, Allowing Work Zone on Both Lanes in Direction 1



(d) Alternative 4.4: Crossover of All Q_1 into One Lane in Opposite Direction, Allowing Work Zone on Both Lanes in Direction 1

Figure 5.3 Geometries of Analyzed Work Zones for Four-Lane Two-Way Highways

5.3 Optimization Model without Considering Time-Cost Tradeoff

In this optimization model, the work zone length L_w is chosen as the decision variable. Work zone cost functions for two-lane two-way highways and multiple-lane two-way highways have been formulated in Chapter 3. The optimal L_w can be obtained by setting the partial derivative of the cost function with respect to equal to zero and then solve for L_w.

5.3.1 Work Zone Optimization for Two-Lane Two-Way Highways

The work zone cost per lane length for two-lane two-way highways can be obtained with the following equations.

$$C_t = \frac{C_T}{L_w \cdot n_w} = \frac{C_M + C_D + C_A}{L_w \cdot n_w} = \frac{z_1 + z_2 L_w n_w + (1 + \frac{\theta_a n_a}{10^8}) v_a t_d}{L_w \cdot n_w} \tag{5-7}$$

$t_d = t_{d1} + t_{d2} + t_{d3} + t_{d23}$

$$t_{d1} = \frac{(z_3 + z_4 L_w)[Q_{1m}(C_w - Q_{1m}) + Q_{2m}(C_w - Q_{2m})]}{V_w (C_w - Q_{1m} - Q_{2m})} L_w + (Q_{1m} + Q_{2m})(z_3 + z_4 L_w)(\frac{L_w}{V_w} - \frac{L_w}{V_0}) \tag{5-8}$$

$$t_{d2} = \xi_1 p_1 Q_1 (z_3 + z_4 L_w)(\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{AB}}{V_{AB}}) + \xi_2 p_2 Q_2 (z_3 + z_4 L_w)(\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,DC}} - \frac{L_{AB}}{V_{AB}}) \tag{5-9}$$

$$t_{d3} = \xi_1 Q_3 (z_3 + z_4 L_w)(\frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{CD}}{V_{CD}}) + \xi_2 Q_4 (z_3 + z_4 L_w)(\frac{L_{CD}}{V^*_{d,DC}} - \frac{L_{CD}}{V_{DC}}) \tag{5-10}$$

$$t_{d23,CD} = 0 \qquad \text{when } Q_3 + P_1 Q_1 \le C \tag{5-11a}$$

$$t_{d23,CD} = \frac{1}{2}(1 + \frac{Q_3 + P_1 Q_1 - C_{CD}}{C_{CD} - Q_3})(Q_3 + P_1 Q_1 - C_{CD})(z_3 + z_4 L_w)^2 \qquad \text{when } Q_3 + P_1 Q_1 > C_{CD} \tag{5-11b}$$

$$t_{d23,DC} = 0 \qquad \text{when } Q_4 + P_2 Q_2 \le C_{DC} \tag{5-12a}$$

$$t_{d23,DC} = \frac{1}{2}(1 + \frac{Q_4 + P_2 Q_2 - C_{DC}}{C_{DC} - Q_4})(Q_4 + P_2 Q_2 - C_{DC})(z_3 + z_4 L_w)^2 \qquad \text{when } Q_4 + P_2 Q_2 > C_{DC} \tag{5-12b}$$

ξ_1 = 1 if the traffic in Direction 1 will be diverted to the Direction 3, otherwise ξ_1=0;

ξ_2 = 1 if the traffic in Direction 2 will be diverted to the Direction 4, otherwise ξ_2=0;

The formulation of optimal work zone length for four alternatives is shown below:

$$L^* = \sqrt{\frac{(P_3 + P_4) z_3 + (P_5 + P_6)(z_3)^2 + z_1}{(P_1 + P_2) z_4 + (P_5 + P_6)(z_4)^2}} \tag{5-13}$$

where,

(1) Alternative 2.1

$$P_1=\frac{v_a[Q_{1m}(C_w-Q_{1m})+Q_{2m}(C_w-Q_{2m})]}{V_w(C_w-Q_{1m}-Q_{2m})} \quad (5\text{-}14a)$$

$$P_2=v_a(Q_{1m}+Q_{2m})(\frac{1}{V_w}-\frac{1}{V_{AB}}) \quad (5\text{-}14b)$$

$$P_3=P_4=P_5=P_6=0 \quad (5\text{-}14c)$$

(2) Alternative 2.2

$$P_1=\frac{v_a[Q_{1m}(C_w-Q_{1m})+Q_{2m}(C_w-Q_{2m})]}{V_w(C_w-Q_{1m}-Q_{2m})} \quad (5\text{-}15a)$$

$$P_2=v_a(Q_{1m}+Q_{2m})(\frac{1}{V_w}-\frac{1}{V_{AB}}) \quad (5\text{-}15b)$$

$$P_3=v_a p_1 Q_1(\frac{L_{AC}}{V_{AC}}+\frac{L_{BD}}{V_{BD}}+\frac{L_{CD}}{V^*_{d,CD}}-\frac{L_{AB}}{V_{AB}})+v_a Q_3(\frac{L_{CD}}{V^*_{d,CD}}-\frac{L_{CD}}{V_{CD}}) \quad (5\text{-}15c)$$

$$P_5=v_a\frac{1}{2}(1+\frac{\max(0,Q_3+P_1Q_1-C_{CD})}{C_{CD}-Q_3})(1+\max(0,Q_3+P_1Q_1-C_{CD})) \quad (5\text{-}15d)$$

$$P4=P6=0 \quad (5\text{-}15e)$$

(3) Alternative 2.3

$$P_1=0 \quad (5\text{-}16a)$$

$$P_2=v_a Q_{2m}(\frac{1}{V_w}-\frac{1}{V_{AB}}) \quad (5\text{-}16b)$$

$$P_3=v_a p_1 Q_1(\frac{L_{AC}}{V_{AC}}+\frac{L_{BD}}{V_{BD}}+\frac{L_{CD}}{V^*_{d,CD}}-\frac{L_{AB}}{V_{AB}})+v_a Q_3(\frac{L_{CD}}{V^*_{d,CD}}-\frac{L_{CD}}{V_{CD}}) \quad (5\text{-}16c)$$

$$P_5=v_a\frac{1}{2}(1+\frac{\max(0,Q_3+P_1Q_1-C_{CD})}{C_{CD}-Q_3})(1+\max(0,Q_3+P_1Q_1-C_{CD})) \quad (5\text{-}16d)$$

$$P_4=P_6=0 \quad (5\text{-}16e)$$

(4) Alternative 2.4

$$P_1=0 \quad (5\text{-}17a)$$

$$P_2=v_a Q_{2m}(\frac{1}{V_w}-\frac{1}{V_{AB}}) \quad (5\text{-}17b)$$

$$P_3=v_a p_1 Q_1(\frac{L_{AC}}{V_{AC}}+\frac{L_{BD}}{V_{BD}}+\frac{L_{CD}}{V^*_{d,CD}}-\frac{L_{AB}}{V_{AB}})+v_a Q_3(\frac{L_{CD}}{V^*_{d,CD}}-\frac{L_{CD}}{V_{CD}}) \quad (5\text{-}17c)$$

$$P_5 = v_a \frac{1}{2}(1+\frac{\max(0, Q_3 + P_1 Q_1 - C_{CD})}{C_{CD} - Q_3})(1+\max(0, Q_3 + P_1 Q_1 - C_{CD})) \quad (5\text{-}17d)$$

$$P_4 = v_a p_2 Q_2 (\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,DC}} - \frac{L_{AB}}{V_{AB}}) + v_a Q_4 (\frac{L_{CD}}{V^*_{d,DC}} - \frac{L_{CD}}{V_{CD}}) \quad (5\text{-}17e)$$

$$P_5 = v_a \frac{1}{2}(1+\frac{\max(0, Q_4 + P_2 Q_2 - C_{DC})}{C_{DC} - Q_4})(1+\max(0, Q_4 + P_2 Q_2 - C_{DC})) \quad (5\text{-}17f)$$

5.3.2 Work Zone Optimization for Multiple-Lane Two-Way Highways

The work zone cost per lane length for two-lane two-way highways can be obtained by the following formulations.

$$C_t = \frac{C_T}{L_w \cdot n_w} = \frac{C_M + C_D + C_A}{L_w \cdot n_w} = \frac{z_1 + z_2 L_w n_w + (1+\frac{\theta_a n_a}{10^8}) v_a t_d}{L_w \cdot n_w} \quad (5\text{-}18)$$

$$t_d = t_{d1} + t_{d2} + t_{d3} + t_{d23} = (t_{d11,queuing} + t_{d11,moving} + t_{d12,queuing} + t_{d12,moving}) + t_{d3} + t_{d23} \quad (5\text{-}19)$$

$$t_{d1i,queuing} = 0 \qquad \text{when } Q_{mi} \le c_{wi} \qquad i=1,2 \quad (5\text{-}20a)$$

$$t_{d1i,queuing} = \frac{1}{2}(1+\frac{Q_{mi} - c_{wi}}{c_{0i} - Q_{mi}})(Q_{mi} - c_{wi})(z_3 + z_4 L_w)^2 \qquad \text{when } Q_{mi} > c_{wi} \qquad i=1,2 \quad (5\text{-}20b)$$

$$t_{d1i,moving} = (\frac{L_w}{V_w} - \frac{L_w}{V_{AB}}) Q_{mi} (z_3 + z_4 L_w) \qquad \text{when } Q_{mi} \le c_{wi} \quad (5\text{-}21a)$$

$$t_{d1i,moving} = (\frac{L_w}{V_w} - \frac{L_w}{V_{AB}}) c_{wi} (z_3 + z_4 L_w) \qquad \text{when } Q_{mi} > c_{wi} \quad (5\text{-}21b)$$

$$t_{d2} = \xi_1 p_1 Q_1 (z_3 + z_4 L_w)(\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{AB}}{V_{AB}}) + \xi_2 p_2 Q_2 (z_3 + z_4 L_w)(\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,DC}} - \frac{L_{AB}}{V_{AB}}) \quad (5\text{-}23)$$

$$t_{d3} = \xi_1 Q_3 (z_3 + z_4 L_w)(\frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{CD}}{V_{CD}}) + \xi_2 Q_4 (z_3 + z_4 L_w)(\frac{L_{CD}}{V^*_{d,DC}} - \frac{L_{CD}}{V_{DC}}) \quad (5\text{-}24)$$

$$t_{d23,CD} = 0 \qquad \text{when } Q_3 + P_1 Q_1 \le C \quad (5\text{-}25a)$$

$$t_{d23,CD} = \frac{1}{2}(1+\frac{Q_3 + P_1 Q_1 - C_{CD}}{C_{CD} - Q_3})(Q_3 + P_1 Q_1 - C_{CD})(z_3 + z_4 L_w)^2 \qquad \text{when } Q_3 + P_1 Q_1 > C_{CD} \quad (5\text{-}25b)$$

$$t_{d23,DC} = 0 \qquad \text{when } Q_4 + P_2 Q_2 \le C_{DC} \quad (5\text{-}26a)$$

$$t_{d23,DC} = \frac{1}{2}(1+\frac{Q_4 + P_2Q_2 - C_{DC}}{C_{DC} - Q_4})(Q_4 + P_2Q_2 - C_{DC})(z_3 + z_4L_w)^2 \qquad \text{when } Q_4 + P_2Q_2 > C_{DC} \qquad (5\text{-}26b)$$

The formulations of optimal work zone length for four alternatives are shown in the following expressions.

$$L^* = \sqrt{\frac{P_3z_3 + (P_1 + P_5)(z_3)^2 + z_1}{P_2z_4n_w + (P_1 + P_5)(z_4n_w)^2}} \qquad (5\text{-}27)$$

where,

(1) Alternative 4.1

$$P_1 = v_a\frac{1}{2}(1+\frac{\max(0, Q_{1m} - C_w)}{C_{AB} - Q_1})(1+\max(0, Q_{1m} - C_w)) \qquad (5\text{-}28a)$$

$$P_2 = v_aQ_{1m}(\frac{1}{V_w} - \frac{1}{V_{AB}}) \qquad (5\text{-}28b)$$

$$P_3 = P_5 = 0 \qquad (5\text{-}28c)$$

(2) Alternative 4.2

$$P_1 = v_a\frac{1}{2}(1+\frac{\max(0, Q_{1m} - C_w)}{C_{AB} - Q_1})(1+\max(0, Q_{1m} - C_w)) \qquad (5\text{-}29a)$$

$$P_2 = v_aQ_{1m}(\frac{1}{V_w} - \frac{1}{V_{AB}}) \qquad (5\text{-}29b)$$

$$P_3 = v_ap_1Q_1(\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{AB}}{V_{AB}}) + v_aQ_3(\frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{CD}}{V_{CD}}) \qquad (5\text{-}29c)$$

$$P_5 = v_a\frac{1}{2}(1+\frac{\max(0, Q_3 + P_1Q_1 - C_{CD})}{C_{CD} - Q_3})(1+\max(0, Q_3 + P_1Q_1 - C_{CD})) \qquad (5\text{-}29d)$$

(3) Alternative 2.3

$$P_1 = 0 \qquad (5\text{-}30a)$$

$$P_2 = 0$$

$$P_3 = v_ap_1Q_1(\frac{L_{AC}}{V_{AC}} + \frac{L_{BD}}{V_{BD}} + \frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{AB}}{V_{AB}}) + v_aQ_3(\frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{CD}}{V_{CD}}) \qquad (5\text{-}30b)$$

$$P_5 = v_a\frac{1}{2}(1+\frac{\max(0, Q_3 + P_1Q_1 - C_{CD})}{C_{CD} - Q_3})(1+\max(0, Q_3 + P_1Q_1 - C_{CD})) \qquad (5\text{-}30c)$$

(4) Alternative 2.4

$$P_1 = v_a\frac{1}{2}(1+\frac{\max(0, Q_{1m} - C_{1w})}{C_{AB} - Q_1})(1+\max(0, Q_{1m} - C_{1w}))$$

$$+\ v_a \frac{1}{2}(1+\frac{\max(0,Q_{2m}-C_{2w})}{C_{AB}-Q_2})(1+\max(0,Q_{2m}-C_{2w})) \quad (5\text{-}31a)$$

$$P_2 = v_a(Q_{1m}+Q_{2m})(\frac{1}{V_w}-\frac{1}{V_{AB}}) \quad (5\text{-}31b)$$

$$P_3=0 \quad (5\text{-}31c)$$

$$P_5=0 \quad (5\text{-}31d)$$

5.4 Optimization Model Considering Time-Cost Tradeoff

The maintenance time and cost can be affected by different resource combinations and construction methods. For congested highway sections, it may be worth spending more on equipment and/or labor in order to significantly reduce the lane closure duration and hence decrease the motorist delays.

Therefore, in this subsection we consider the time and cost tradeoffs in planning the road maintenance projects. The work zone length and the work rate parameter are simultaneously optimized to minimize the work zone total cost per completed lane length work.

5.4.1 Proposed Relation of Time and Cost

In previous studies without considering time and cost tradeoff (Chen and Schonfeld, 2005), it is assumed that the cost and duration functions are both linear. The total construction cost function for a given work zone is defined as $C_M=z_1+z_2L_w$; the construction cost per length C_m is then the total cost C_M divided by the length L_w:

$$C_m = \frac{z_1}{L_w} + z_2 \quad (5\text{-}32)$$

where z_1 is the fixed set-up cost per zone and z_2 is the variable cost per lane length. The z_2 value would depend on worker numbers and skills, construction methods, materials, equipment productivities, etc.

The work duration for a work zone of length L_w is:

$$D_w = z_3 + z_4 L_w \quad (5\text{-}33)$$

where z_3 is a fixed setup time per zone and z4 is the variable work duration needed per lane length.

In introducing tradeoffs between work duration and cost, we note that any point on a time-cost tradeoff function (such as A, B, C in Figure 5.4) uniquely determines both a variable cost parameter z_2 (in \$/lane-length) and a variable time parameter z_4 (in hrs/lane-length). For mathematical convenience we choose to use z_2 rather than z_4 as one optimizable "work rate" parameter. Of course, optimizing either one of them also optimizes the other and hence resolves the time-cost tradeoff.

For reasonable simplification and applicability we assume here that the function relating variable cost to variable time is a shifted hyperbolic one, which is convex and continuously differentiable. The functional relation between variable cost and variable time per lane per length is defined in a hyperbolic form.

$$(z_2 - c_v)(z_4 - t_v) = k_1 \tag{5-34}$$

In Eq.(3), k_1 is the trade-off coefficient that can be estimated through empirical data; it can represent various types of construction activities, e.g. grinding, paving, or reconstruction; t_v is the minimum variable duration required per lane per length; c_v is the minimum variable cost per lane per length.

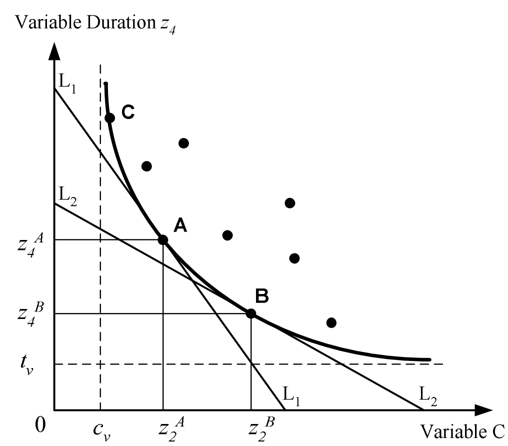


Figure 5.4 Time-Cost Tradeoff Relation

Thus, the variable duration z_4 can then be represented as a function of the variable cost z_2.

$$z_4 = t_v + \frac{k_1}{z_2 - c_v} \tag{5-35}$$

The choice of z_2, which can also be treated as the "work rate" parameter, depends on how one evaluates the time value of the construction work. For congested facilities or congested periods when the user delay may increase drastically, construction agencies may prefer to employ more equipment and manpower to accelerate the work and reduce traffic impacts. In Figure 5.4, the lines L_1 and L_2 represent two different ratios of time and cost. As the cost of user delays increases, the work should be done faster (but more expensively). Thus, L_1 has a lower time value (cost per time unit) than L_2, and the contact points A and B indicate the optimal tradeoff combination of (z_2, z_4) for the time value lines L_1 and L_2, respectively.

5.4.2 Optimization Model Formulation

The total cost per lane-length, which constitutes our objective function, has been formulated in Eqs.(5-3) to (5-31).

The delay component of the total cost and hence the structure of the objective function depends significantly on whether the inflow exceeds the capacity. Therefore, two different objective functions are used to optimize the system, depending on whether inflows Q_{1m} exceed the work zone capacity C_w.

Applying $z_4 = t_v + \frac{k_1}{z_2 - c_v}$,

for $Q_{1m} \le C_w$, we have

$$C_T = \frac{z_1}{L_w} + z_2 + P_1\left\{Q_{1m}P_2\left[z_3 + \left(t_v + \frac{k_1}{z_2 - c_v}\right)L\right] + P_3\left[\frac{z_3}{L_w} + \left(t_v + \frac{k_1}{z_2 - c_v}\right)\right]\right\} + P_4\left[\frac{z_3}{L_w} + \left(t_v + \frac{k_1}{z_2 - c_v}\right)\right] \tag{5-36a}$$

for $Q_{1m} > C_w$, we have

$$C_T = \frac{z_1}{L_w} + z_2 + P_1\left\{\begin{matrix} \frac{P_5}{L_w}\left[z_3 + \left(t_v + \frac{k_1}{z_2 - c_v}\right)L_w\right]^2 + c_w P_2\left[z_3 + \left(t_v + \frac{k_1}{z_2 - c_v}\right)L_w\right] \\ + P_3\left[\frac{z_3}{L_w} + \left(t_v + \frac{k_1}{z_2 - c_v}\right)\right] \end{matrix}\right\} + P_4\left[\frac{z_3}{L_w} + \left(t_v + \frac{k_1}{z_2 - c_v}\right)\right] \tag{5-36b}$$

where

$$P_1 = \left(v_d + \frac{n_a v_a}{10^8}\right) \tag{5-37a}$$

$$P_2 = \left(\frac{1}{V_w} - \frac{1}{V_{AB}}\right) \tag{5-37b}$$

$$P_3 = p_1 Q_1\left(\frac{L_{AC}}{V_{AC}} + \frac{L_{DB}}{V_{DB}} + \frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{AB}}{V_{AB}}\right) \tag{5-37c}$$

$$P_4 = v_d Q_3\left(\frac{L_{CD}}{V^*_{d,CD}} - \frac{L_{CD}}{V_{CD}}\right) \tag{5-37d}$$

$$P_5 = \frac{(Q_{1m} - C_w)}{2}\left(1 + \frac{Q_{1m} - C_w}{C_{AB} - Q_{1m}}\right) \tag{5-37e}$$

Here P_1, P_2, P_3, and P_4 are defined for convenience of expression.

It can be determined by inspection that Eq. (5-36a) and Eq. (5-36b) are convex with respect to (L_w, z_2). The convexity ensures that the solution satisfying the optimality conditions is globally optimal. In the following analysis, we derive and analyze the optimality conditions based on the above two formulations.

(1) First-order Conditions

The first-order conditions for $Q_{1m} \le C_w$ are shown below.

$$\frac{\partial C_T}{\partial L_w} = 0 = P_1 P_2 Q_{1m}\left(t_v + \frac{k_1}{z_2 - c_v}\right) - \left(z_1 + P_1 P_3 z_3 + P_4 z_3\right) L_w^{-2} \tag{5-38}$$

$$\frac{\partial C_T}{\partial z_2} = 0 = 1 - k_1\left(P_1 P_2 Q_{1N} L + P_1 P_3 + P_4\right)\left(z_2 - c_v\right)^{-2} \tag{5-39}$$

For $Q_{1m} > C_w$, the first-order conditions are:

$$\frac{\partial C_T}{\partial L_w} = 0 = P_1 P_2 c_w\left(t_v + \frac{k_1}{z_2 - c_v}\right) + P_1 P_5\left(t_v + \frac{k_1}{z_2 - c_v}\right)^2 - \left(z_1 + P_1 P_3 z_3 + P_4 z_3 + P_1 P_5 z_3^{\,2}\right) L_w^{-2} \tag{5-40}$$

$$\frac{\partial C_T}{\partial z_2} = 0 = 1 - k_1\left(2P_1 P_5 z_3 + P_1 P_2 C_w L_w + P_1 P_3 + P_4 + 2P_1 P_5 t_v L_w\right)\left(z_2 - c_v\right)^{-2} - 2P_1 P_5 k_1^{\,2} L_w \left(z_2 - c_v\right)^{-3} \tag{5-41}$$

By solving each of the equation pairs above we can obtain the applicable optimal combination of (L_w, z_2) that minimizes total system.

(2) Second-order Conditions: Characteristics of Hessian

To satisfy the second order conditions for the minimum, we must also check that at the optimal values L_w* and z_2* the Hessian matrix of the cost function C_T must be positive definite (Nash and Sofer, 1996). The derived Hessian matrix is symmetric and shown below.

$$H = \begin{bmatrix} \dfrac{\partial^2 C_T}{\partial L_w^{\,2}} & \dfrac{\partial^2 C_T}{\partial L_w \partial z_2} \\ \dfrac{\partial^2 C_T}{\partial z_2 \partial L_w} & \dfrac{\partial^2 C_T}{\partial z_2^{\,2}} \end{bmatrix} \tag{5-42}$$

For Q_{1m} not exceeding the capacity

$$\frac{\partial^2 C_T}{\partial L_w^{\,2}} = 2\left(z_1 + P_1 P_3 z_3 + P_4 z_3\right) L_w^{-3} \tag{5-43}$$

$$\frac{\partial^2 C_T}{\partial z_2^{\,2}} = 2k_1\left(P_1 P_2 Q_{1m} L_w + P_1 P_3 + P_4\right)\left(z_2 - c_v\right)^{-3} \tag{5-44}$$

$$\frac{\partial^2 C_T}{\partial L_w \partial z_2} = \frac{\partial^2 C_T}{\partial z_2 \partial L_w} = -P_1 P_2 Q_{1m} k_1 (z_2 - c_v)^{-2} \tag{5-45}$$

For Q_{1m} exceeding the capacity

$$\frac{\partial^2 C_T}{\partial L_w^{\,2}} = 2\left(z_1 + P_1 P_3 z_3 + P_4 z_3 + P_1 P_5 z_3^{\,2}\right) L_w^{-3} \tag{5-46}$$

$$\frac{\partial^2 C_T}{\partial {z_2}^2} = 2\left(2P_1P_5k_1z_3 + P_1P_2c_wk_1L_w + P_1P_3k_1 + P_4k_1 + 2P_1P_5k_1t_vL_w\right)\left(z_2 - c_v\right)^{-3} + 6P_1P_5{k_1}^2L_w\left(z_2 - c_v\right)^{-4} \quad (5\text{-}47)$$

$$\frac{\partial^2 C_T}{\partial L_w \partial z_2} = \frac{\partial^2 C_T}{\partial z_2 \partial L_w} = -\left(P_1P_2C_wk_1 + 2P_1P_5k_1t_v\right)\left(z_2 - c_v\right)^{-2} - 2P_1P_5k_1^2\left(z_2 - c_v\right)^{-3} \quad (5\text{-}48)$$

For any reasonable numerical values (i.e., positive but not impractically high) that are substituted into the above equations, we can confirm the positive definite property of the Hessian, and hence the global optimality of the solution.

5.5 Numerical Example

The numerical values chosen for this analysis are based on previous studies and rough estimates. They illustrate the model's capabilities, but would be replaced by more precise and pertinent values for practical applications. After applying the baseline values to the optimality conditions, we can obtain the minimum cost solution. The baseline values of all variables used in the example and the numerical results are summarized in Table 5.1 and Table 5.2, respectively.

Sensitivity Analysis

To illustrate the effects of the tradeoff between time and cost, at a given diversion fraction $p = 0.3$, we analyze the sensitivity of the resulting optimal work zone length, variable cost, and total system cost to the inflow Q_1. The results are summarized in Table 5.2 and shown in Figure 5.5.

(1) Inflow vs. Total System Cost

As expected and shown in Figure 5.6, the minimized total cost increases as Q_1 increases. After the undiverted flow Q_{1N} exceeds the work zone capacity c_w, the total cost increases rapidly as the user delay cost increases rapidly and accounts for an increasing fraction of the total system cost. A larger z_2 is then optimized to reduce the work duration.

Table 5.1 Notation and Baseline Numerical Inputs

Variable	Description	Baseline Values
C_a	Average accident cost per lane-kilometer ($/lane-km)	
C_M	Maintenance cost ($/lane-km)	
c_o	Maximum discharge rate without work zone (vph)	2,600 vph
C_q	Queue delay cost ($/lane-km)	
C_T	Average total cost ($/lane-km)	
C_u	Average user cost ($/lane-km)	
C_v	Moving delay cost ($/lane-km)	
c_w	Maximum discharge rate along work zone (vph)	1,200 vph
D	Total work duration for work zone length L (hr)	
d	Average work duration ($/lane-km)	
L	Work zone length (km)	
L_d	Detour length (km), which is $L_{d1}+L_{d2}+L_{d3}$	
L_{d1}	Length of first detour segment (km)	0.5 km
L_{d2}	Length of second detour segment (km)	5 km
L_{d3}	Length of third detour segment (km)	0.5 km
L_t	Length from A to B (km), which is $L+L_1+L_2$	5 km
K_j	Jam density (veh/lane-km)	200 veh/lane·km
n_a	Number of accidents per 100 million vehicle hours (number of acc/ 100 mvh)	40 acc/100mvh
p	The fraction of flow in Direction 1 that diverts to alternate route	0.3
Q_1	Hourly flow rate in Direction 1 (veh/hr)	1600
Q_{1N}	The undiverted flow in Direction 1 (veh/hr)	
Q_{1D}	The diverted flow in Direction 1 (veh/hr)	
Q_3	Hourly flow rate in Direction 3 (veh/hr)	500 vph
t_d	Queue dissipation time (hr)	
t_m	Moving delay (veh-hr)	
t_q	Queueing delay (veh-hr)	
V_a	The approaching speed on original road without work zone (km/hr)	
V_b	The speed of diverted traffic Q_{1D} on the segments of L_{d1} and L_{d3},	
V_f	Free flow speed along AB and detour (km/h)	80 km/hr
V_w	Work zone speed limit (km/hr)	50 km/hr
V_d^*	Detour speed in Direction 3 affected by diverted traffic from Direction 1 (km/hr)	
V_{d0}	Original speed on L_{d2} unaffected by Q_1 (km/hr)	
v_a	Average accident cost ($/accident)	142,000 $/acc
v_d	Value of user time ($/veh·hr)	12 $/veh·hr
z_1	Fixed setup cost ($/zone)	1,000 $/zone
z_2	Average maintenance cost per additional lane-km ($/lane-km)	
z_3	Fixed setup time (hr/zone)	2 hr/zone
z_4	Average maintenance time (hr/lane-km)	
k_1	Tradeoff coefficient, depending on the types of construction activities; ($·hr/(lane-km)^2)	12,000$·hr/(lane-km)^2
t_v	Minimum variable time per lane-kilometer (hr/lane-km)	3 hr/lane-km
c_v	Minimum variable cost per lane-kilometer ($/lane-km)	50,000 $/lane-km

Table 5.2 Numerical Results with *p*=0.3 and k_1=12,000

Inflow from Direction 1 (vph) Q_1	Optimal Work Zone Length (km) L^*	Optimal Variable Cost ($/lane-km) z_2^*	Minimum Total System Cost ($/lane-km) C_T^*	Variable Duration (hr/lane-km) z_4	Total User Delay Cost ($/lane-km) $C_u^N + C_u^D$
400	1.5885	50,860.2340	52,608.8307	16.9497	1,059.0870
500	1.4841	50,936.2388	52,859.0097	15.8172	1,178.0461
600	1.4051	51,002.5091	53,081.3353	14.9700	1,285.5335
700	1.3424	51,061.2600	53,282.0246	14.3073	1,383.7223
800	1.2910	51,113.8882	53,464.9702	13.7731	1,473.9303
900	1.2477	51,161.3274	53,632.7522	13.3330	1,556.9873
1000	1.2104	51,204.2253	53,787.1418	12.9649	1,633.4180
1100	1.1778	51,243.0407	53,929.3765	12.6537	1,703.5407
1200	1.1489	51,278.0998	54,060.3209	12.3889	1,767.5246
1300	1.1229	51,309.6306	54,180.5646	12.1629	1,825.4236
1400	1.0992	51,337.7853	54,290.4850	11.9700	1,877.1963
1500	1.0774	51,362.6531	54,390.2868	11.8064	1,922.7180
1600	1.0570	51,384.2691	54,480.0269	11.6688	1,961.7859
1700	1.0379	51,402.6186	54,559.6301	11.5554	1,994.1204
1710	1.0361	51,404.2718	54,567.0263	11.5454	1,996.9703
1715	0.9558	51,654.2428	55,078.6041	10.2541	2,193.5704
1720	0.7509	52,491.5673	57,258.8632	7.8162	3,277.7873
1750	0.5701	54,538.3786	64,704.1207	5.6441	8,268.8357
1800	0.5353	56,546.2369	74,223.5376	4.8331	15,667.6681
1900	0.5344	59,475.5856	91,782.6908	4.2664	30,288.5455
2000	0.5432	61,909.3514	109,647.4934	4.0076	45,740.4755
2100	0.5522	64,149.5342	128,712.4138	3.8481	62,583.9686
2200	0.5605	66,316.7369	149,545.6230	3.7354	81,263.7682
2300	0.5680	68,481.3648	172,699.0909	3.6493	102,261.6614
2400	0.5748	70,695.8989	198,811.3888	3.5798	126,164.2093
2500	0.5811	73,007.5667	228,685.2741	3.5216	153,727.2318
2600	0.5869	75,465.5330	263,381.1675	3.4712	185,962.0189

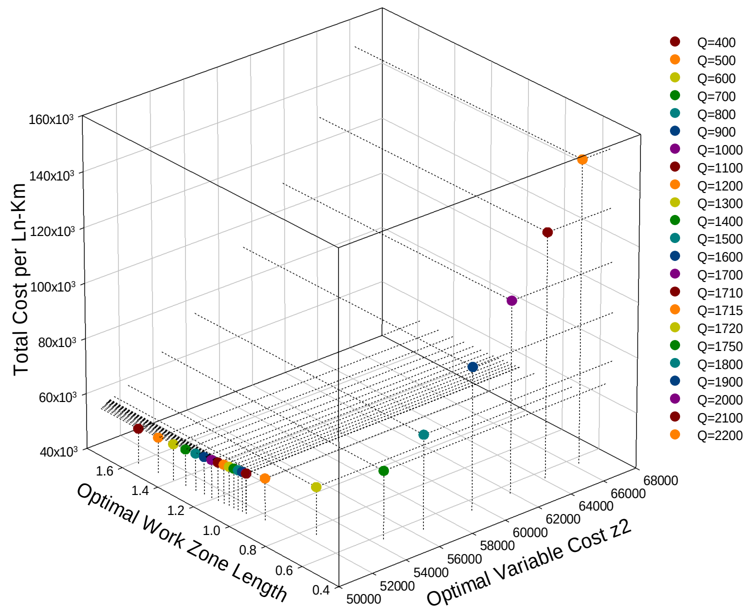


Figure 5.5 Optimization Results with p=0.3 and k1=12,000

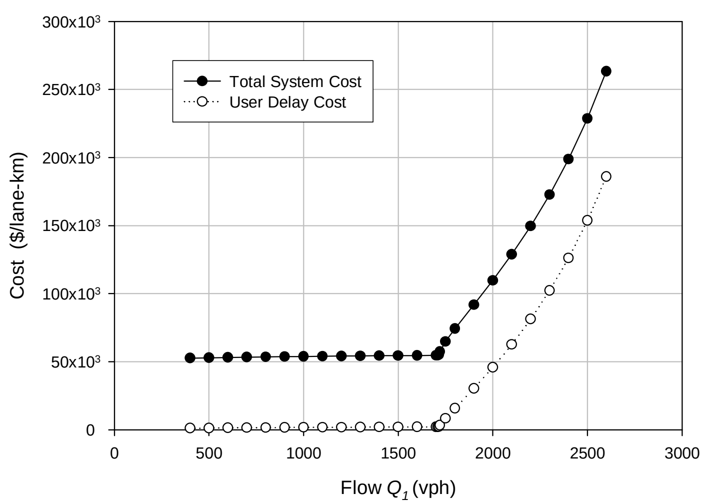


Figure 5.6 Minimum Total Cost and User Delay Cost for Various Inflows

(2) Inflow vs. Variable Cost

Figure 5.7 shows z_2* and C_T* for various flow conditions. When $Q_{1N} < c_w$, the optimized variable cost z_2* does not increase significantly; z_2* increases rapidly to reduce the work duration after $Q_{1N} > c_w$.

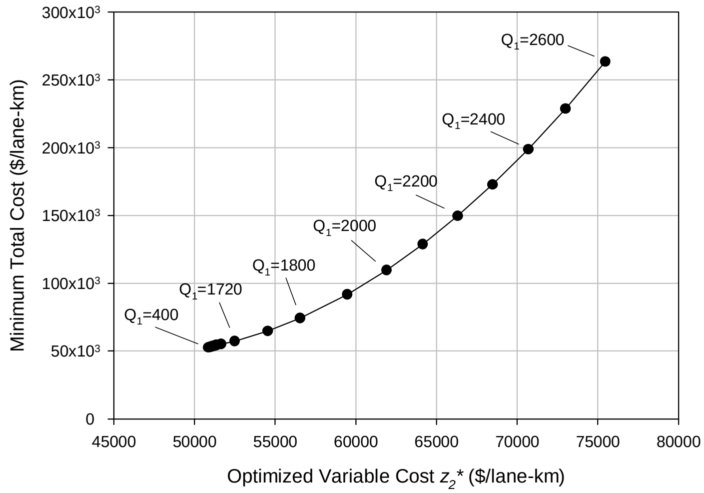


Figure 5.7 Minimum Total Cost versus Optimized Variable Cost for Various Inflows

(3) Variable Cost vs. Variable Time

The effect of inflow Q_1 on optimal combinations of z_2 and z_4 is shown in Figure 5.8. When the $Q_{1N} < c_w$, the optimal combinations of (z_2, z_4) are located in the upper part of the hyperbolic function. As Q_1 increases, the slightly increasing z_2 would significantly decrease the variable duration z_4. When $Q_{1N} > c_w$, the optimal combinations of (z_2, z_4) move toward the lower part of the hyperbolic function, where z_2 increases faster than z_4 decreases. Thus, higher user delay cost leads the optimization to faster but more expensive construction practices. Reducing construction time to reduce traffic delay justifies higher construction costs.

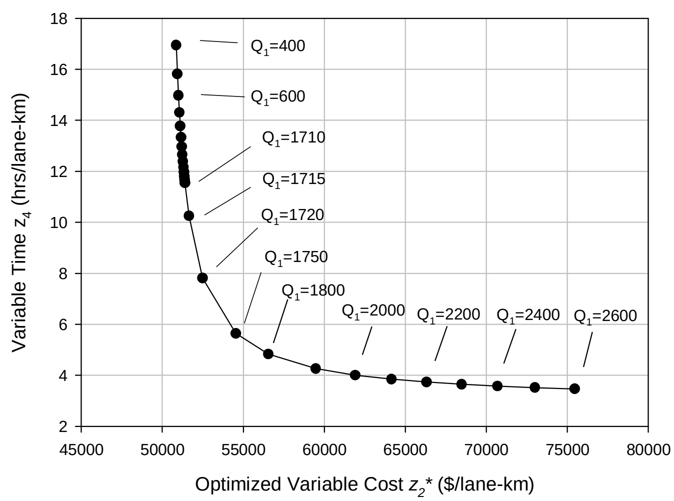


Figure 5.8 Variable Time versus Optimized Variable Cost for Various Inflows

(4) Work Zone Length vs. Inflow and Work Zone Length vs. Total Cost

In Figures 5.9 and 5.10, when $Q_{1N} < c_w$, the optimized length L^* changes much more than z_2^* as Q_{1N} increases gradually. However, L^* drops suddenly after Q_{1N} exceeds the work zone capacity, although it may later increase slowly as the inflow keeps increasing.

Previous studies indicate that the optimized length L^* decreases as the inflow Q_1 increases (Schonfeld et al., 1999; Chien et al., 2001; Chen et al., 2005). Those results differ from our present ones when the work zone capacity is exceeded because we additionally optimize the time-cost tradeoff combination. In Figure 5.8, for $Q_{1N} > c_w$ we compare how the optimal solutions change as the setup cost z_1 changes from \$1,000 to \$5,000 per zone. The optimized work zone length L^* decreases as Q_1 increases for both $z_1 = 5{,}000$ and $z_1 = 1{,}000$. However, as Q_1 increases after $Q_{1N} > c_w$, the optimal length increases when $z_1 = 1{,}000$, but it decreases when $z_1 = 5{,}000$. Thus, through the joint optimization of L and z_2, L^* may either increase or decrease as Q_1 increases, depending on values of input parameters such as zone setup cost.

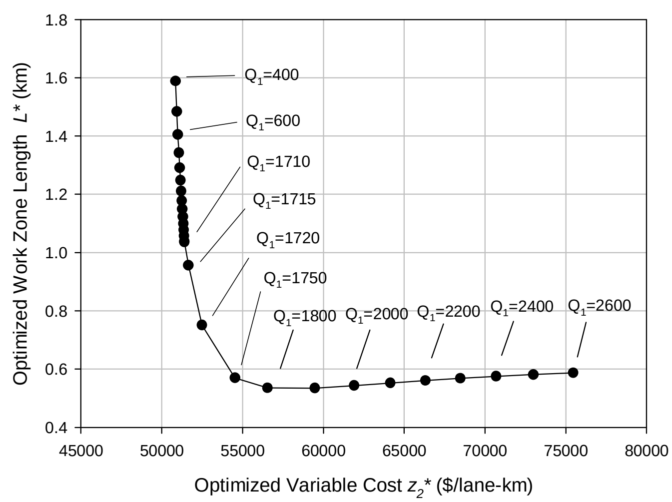


Figure 5.9 Optimized Length versus Optimized Variable Cost for Various Inflows

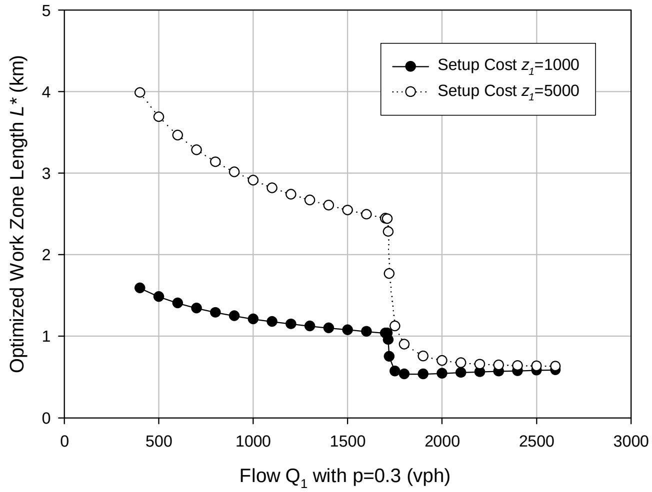


Figure 5.10 Optimized Length for Various Inflows

(5) Determination of Diversion Fraction p

The relation between diversion fraction p and inflows Q_1 is shown in Figure 5.11.

Each line indicates, for various p values, how the minimum total cost C_T* increases as Q_1 increases. The kinks in C_T* curves occur at the critical points when the undiverted flow Q_1 exceeds the work zone capacity c_w and moving delay increases drastically. The optimal diversion fraction can be found along the lowest envelope of the curves in Figure 5.11. The fraction p is then no longer a predetermined parameter and can be optimized according to the traffic conditions.

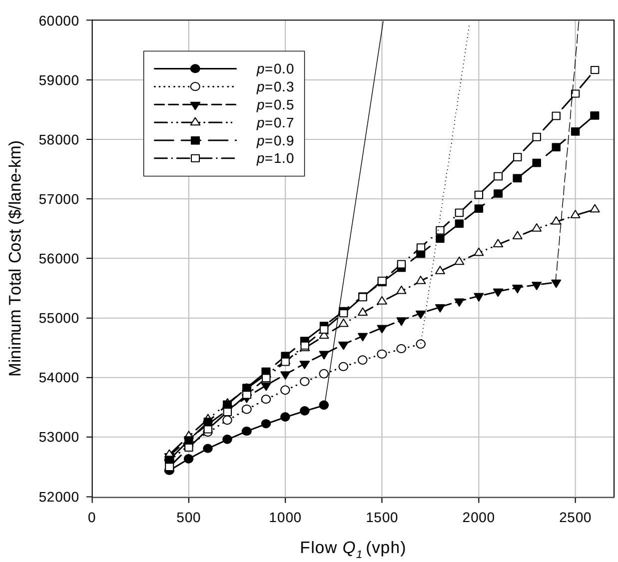


Figure 5.11 Inflows versus Total Cost under Various Diversion Fraction

To sum up, the sensitivity analysis shows that the optimal work rate parameter z_2* increases as the flow through the work zone increases. When a queue occurs, the value of z_2* increases rapidly to minimize the total cost. Even when z4* does not decrease faster than z_2*, employing more expensive but faster construction method can help minimize the total cost.

Unlike in previous studies, when the time-cost tradeoff is incorporated in the optimization, the optimal zone length L* does not always decrease as the inflow Q_1 increases; that length is jointly optimized with z_2 based on several input parameters.

The diversion fraction represents another dimension of work zone optimization. The optimized fraction p* can be obtained by analyzing the envelope of the lowest total cost curves in Figure 5.10. Thus, given the flow conditions, the optimal combination of length, variable cost, and the diversion fraction can be determined.

5.6 Summary

In this chapter, two work zone optimization models for steady traffic inflows are formulated with the objective to minimize the work zone cost per lane length, using the classical optimization method of differential calculus. In one model, the work zone length is used as the decision variable. In another optimization model, the work zone length and work rate, which represents a tradeoff relation between work time and cost, are jointly optimized.

Chapter 6 Two-Stage Modified Simulated Annealing Optimization Algorithm

When no simple formulation is available for the objective function of an optimization problem, a good optimization method is necessary to search in the solution space and reach a good solution quickly, without excessive memory requirements. In this chapter, an optimization algorithm called two-stage modified simulated annealing (2SA) is developed to solve the work zone optimization problem, whose objective function is obtained from the analytic method for time-dependent traffic inflows or the simulation method. This chapter introduces the concept and the procedure of the algorithm. Its application will be presented in the following three chapters.

6.1 Introduction

Simulated annealing (SA) is a stochastic computational technique derived from statistical mechanics for finding near globally optimum solutions to large optimization problems. The SA algorithm exploits the analogy between annealing solids and solving combinatorial optimization problems. This neighborhood search algorithm attempts to avoid being trapped in a local extreme by sometimes moving in locally worse direction. Since SA rapidly modifies small-scale structure within the model, it often finds high quality candidate solutions in doing refined searches inside prominent regions.

However, when solving realistic problems with a large number of parameters and a great complexity, a great deal of work may be required to reach this neighborhood. For example, when the construction work is allowed in a long-term duration such as a whole week, there may exist several local optimal solutions such as 8-hour off-peak daytime windows, 10-hour nighttime windows, 30-hour weekend windows. Suppose 10-hour nighttime windows are the optimal solution. It may take much time for SA to jump out of other local optima and get close to the optimal solution, especially when the initial solution is far away.

To overcome this limitation, we develop a two-stage modified simulated annealing algorithm to solve the work zone optimization problem, where numerous local optima are likely to occur.

There are two-stages in this newly-developed algorithm.

The first stage is initial optimization. In this step, a population-based search procedure in combination with the annealing technique, is employed to obtain an initially optimized solution after a widely search in the relatively large solution space.

The second stage is refined optimization. In this step, a traditional simulated annealing (SA) algorithm is applied. We seek to use this neighborhood search algorithm to find high quality candidate solutions in doing refined searches inside prominent regions provided by the first stage.

This two-stage modified SA algorithm will be applied in the work zone optimization based on analytic method for time-dependent traffic inflows, the work zone optimization based on simulation method and the work zone optimization through hybrid method.

6.2 General Solution Search Procedure

Like all other direct search methods, the two-stage modified SA algorithm is iterative in nature. It starts from a group of initial trial solutions and proceed toward the minimum point in a sequential manner. The major components included in the optimization algorithm are: (1) Solution generation; (3) Solution Modification; (3) Solution evaluation; and (4) Search procedure.

The details of the first three modules depend on the specific optimization problem to be solved. Here we focus on the design of the search procedure.

Before we describe the search procedure design, let's formulate a general statement of an optimization problem. Consider a general function minimization problem, with function $f : \overrightarrow{X} \rightarrow R^D$,

$$J = \min f(\overrightarrow{X}) \tag{6-1}$$

$$\overrightarrow{X} = \{x_1, x_2, ..., x_D\} \tag{6-2}$$

$$x_i \in [l_i, u_i] \tag{6-3}$$

where $\overrightarrow{X}$ is the set of all decision variables, D is the number of the decision variables x_i, f is the objective function and l_i and u_i are the lower and upper bounds, respectively. The goal for this problem is to search for the global minimum f^*, where $\overrightarrow{X}^*$ is the minimum location.

6.2.1 The First Stage - Initial Optimization

In the first stage, a population-based simulated annealing (PBSA) algorithm is developed. The framework of the first stage is shown in Figure 6.1.

STEP1:

Generate the first-generation population $\overrightarrow{X}(N)$, where N is the population size. In the solution generation process, all the variables x_i should satisfy the lower and upper bound constraints.

Evaluate each solution $\overrightarrow{X}(i)$ and then obtain objective function values $C(i) = f(\overrightarrow{X}(i))$. Record the best solution in this generation X^* and C^*.

Set the values of initial temperature T_0 , stop temperature T_f. and step size of the temperature $\triangle T$. Set $T_J = T_0$.

STEP2:

As long as the stopping criterion ($T \leqslant T_f$) is not satisfied, perform the sub-steps.

Step 2.1: Modify of the solution $\overrightarrow{X}(i)$ in the current population to obtain neighboring solutions $\overrightarrow{X_m}(i)$. A check procedure is used to ensure the neighboring solution satisfy the constraints.

Step 2.2: Compute the objective function value and the difference between the new and best recorded value for every solution $\Delta C(i) = C(\overline{X_m}(i)) - C^*$. If $\Delta C(i) < 0$, go to step 2.4. Otherwise, go to step 2.3.

Step 2.3: ($\Delta C(i) > 0$) Select a random variable $\alpha \in U(0,1)$.

If $\alpha < \Pr ob(\Delta C(i)) = \exp(-\Delta C(i)/T_j)$, then go to step 2.4. If $\alpha > \Pr ob(\Delta C(i))$, then reject this new solution and add the previous solution into the next generation. Skip step 2.4.

Step 2.4: ($\Delta C(i) < 0$ or $\alpha < \Pr ob(\Delta C(i))$) Accept the new solution $\overrightarrow{X_m}(i)$ and the new value $C(i) = f(\overline{X_m}(i))$. Add the new solution into the next generation.

Step 2.5: Record the best solution in the new generation X^* and the corresponding C^* if they are changed.

Step 2.6: If $i < N$, $i = i + 1$, go to step 2.1. For each candidate solution in the current generation, repeat the above steps.

Else if i=N, then update $T_J = T_J - \Delta T$, $J = J+1$, $K=1$, and go to step 2.1.

STEP3:

Output the best solution ever found. Pass this solution to the next stage.

6.2.2 The Second Stage- Refined Optimization

In the second stage, the classical simulated annealing (SA) algorithm is used. The framework of the first stage is shown in Figure 6.2.

STEP1:

Use the optimized solution and its objective function value obtained from the first stage as the initial solution $\overrightarrow{X}$ and its objective function value $C=f(\overrightarrow{X})$ in this stage. Record the best solution $X^*=\overrightarrow{X}$ and $C^*=f(\overrightarrow{X})$.

Set the values of initial temperature T_0 and stop temperature T_f. and step size of the temperature $\triangle T$. Choose a repetition factor K_{max}. Set $T= T_0$, $K=0$.

STEP2:

As long as the stopping criterion （$T \leqslant T_f$） is not satisfied, perform the sub-steps.

Step 2.1: Modify the solutions in the current population to obtain neighboring solutions $\overrightarrow{X_m}$. A check procedure is used to ensure the neighboring solution satisfy the constraints.

Step 2.2: Compute the objective function value and the difference between the new and best recorded value for every solution $\Delta C = f(\overrightarrow{X_m}) - C^*$. If $\Delta C < 0$, go to Step 2.4; Otherwise, go to step 2.3.

Step 2.3: ($\Delta C > 0$)   Select a random variable $\alpha \in U(0,1)$ If $\alpha < \Pr ob(\Delta C) = \exp(-\Delta C / T_j)$, then go to step 2.4. If $\alpha > \Pr ob(\Delta C)$, then reject this new solution and add the previous solutions into the next generation. Skip step 2.4.

Step 2.4: ($\Delta C < 0$ or $\alpha < \Pr ob(\Delta C)$) Accept the new solution $\overrightarrow{X_m}$ and the new value $C_m = f(\overrightarrow{X}_m)$. Replace the $\overrightarrow{X}$ and C with $\overrightarrow{X_m}$ and C_m.

Step 2.5: Update the best solution X^* and the corresponding C^* if they are changed.

Step 2.6: If $K < K_{\max}$, then $K \leftarrow K+1$ and go to step 2.1.

Else if $K = K_{\max}$, then reduce T_J , $J = J+1$, $K=1$, and go to step 2.1.

STEP3:

Output the best solution ever found as the final solution.

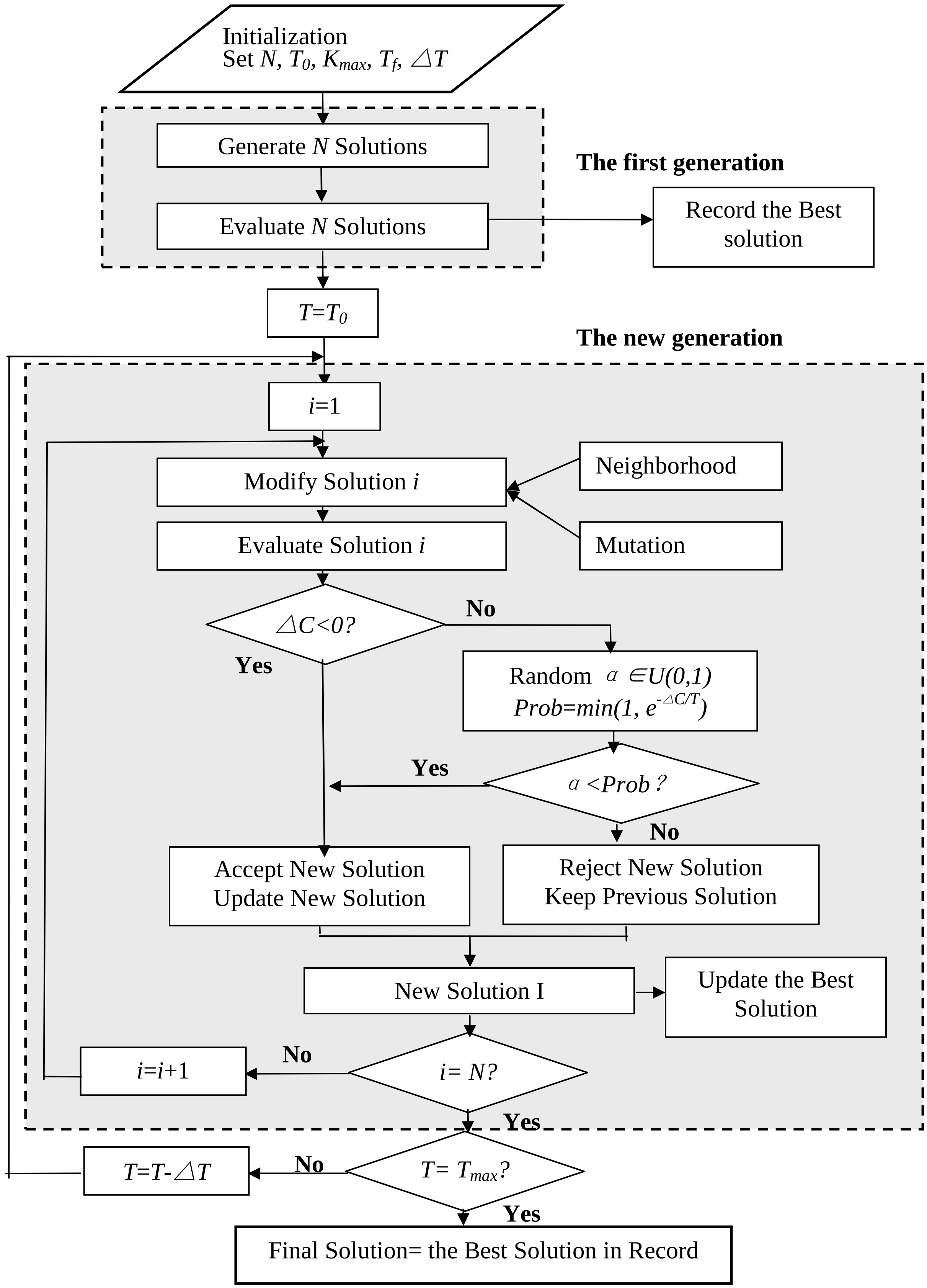


Figure 6.1 Framework of the First Stage (PBSA)

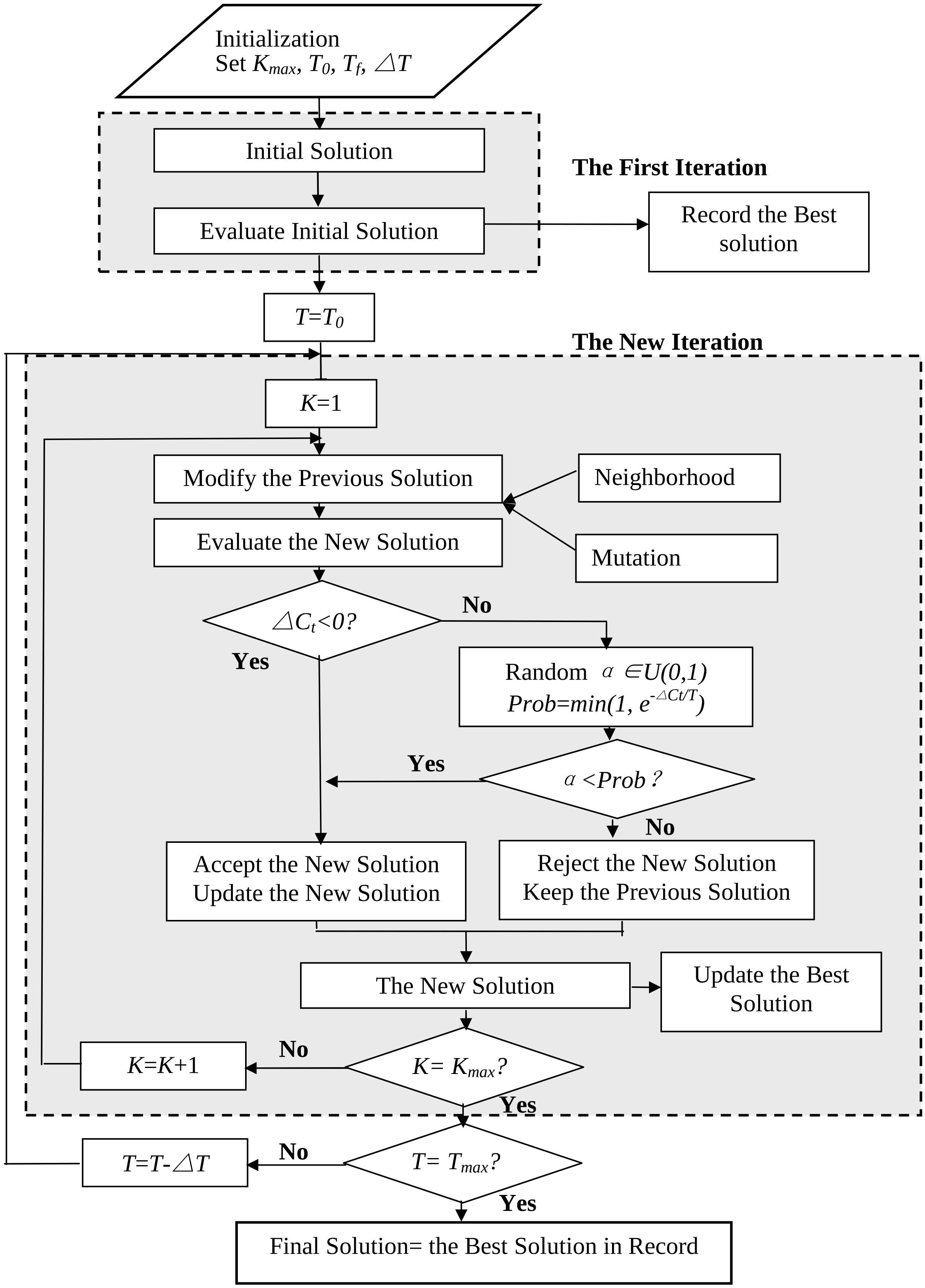


Figure 6.2 Framework of the Second Stage (SA)

6.3 Comparison of PBSA and SA Algorithm

The difference between the proposed two-stage Modified SA (TMSA) approach and the traditional SA is the additional initial optimization stage, in which PBSA algorithm is applied. Therefore, we compare PBSA with SA to test its performance in this subsection.

We perform an experiment on a function with multiple optima. The function is generated by the following equation:

$$f(x_1, x_2) = x_1^{\,2} + 2x_2^{\,2} - 0.3\cos(3\pi x_1) - 0.4\cos(4\pi x_2) + 0.7 \qquad (6\text{-}4)$$

This function was examined by Bohachevsky et al (1986). It has numerous local minima and a global minimum at the origin ($x_1=0$, $x_2=0$). This function is illustrated in Figure 6.3 with $x_1, x_2 \in [-1,1]$.

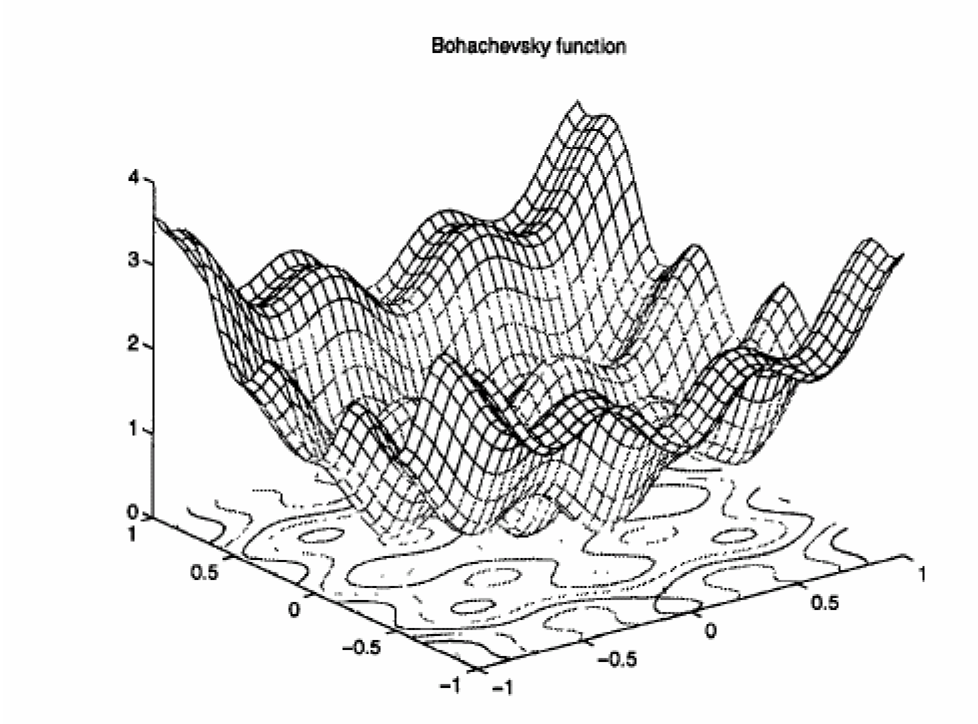


Figure 6.3 The Bohachevsky Function

PBSA and SA are applied to find the global optimum of the Bohachevsky function. In both algorithms, solutions are generated, modified and evaluated in the same way. The initial solutions are generated randomly in the region of [-1,1]. The initial temperature T_0= 0.5, the stop temperature T_f =0 and the step size of temperature change $\triangle T$ =0.01 are used in both

PBSA and SA. To equalize the number of solution evaluations in both algorithms, the population size N in PBSA is set to be the same as the repetition factor K_{max} in SA.

This experiment is run on a Pentium® 4 3.60 GHz PC with 0.99 GB of RAM.

(1) Comparison of convergence within the same number of generation

Setting N=$Kmax$=200, we compared 20 independent runs of PBSA and SA. The results are shown in Table 6.1. The convergences of the two algorithms are displayed in Figure 6.4. We can see the PBSA dominates SA with a higher quality optimized result, a lower standard deviation and a shorter running time. The reason is that SA is a neighborhood search method. It may be slow and inefficient in searching a large solution space with multiple local optima. It may take much time and effort for SA to jump from a local optimum to another one. PBSA uses a population of candidate solutions and a wider solution space can be searched. Besides, there will be more opportunities for PBSA to avoid getting stuck in a local optimum.

(2) Comparison of performance with increasing population size

Varying the value of N=K_{max} ranging from 20 to 250, we intend to check the impact of population size/ repetition factor on the performance of the two algorithms. Table 6.2 shows the statistics analysis of the optimized results for 20 independent runs for PBSA and SA. The changes of the average optimized results with respect to the population size / repetition factor are shown in Figure 6.5. We see that PBSA is more sensitive to the population size, as we expected. The larger the population size, the better the optimized result is and the run-to-run variance is also decreasing. This is consistent with prior expectations. However, SA is not sensitive to the repetition factor.

Table 6.1 Comparison of Convergence (20 runs)

	PBSA	SA
Number of Generations	51	51
Number of Evaluations	10000	10000
Obj. Value (mean)	0.000308	0.036122
Obj. Value (st. dev)	0.000429	0.026006
Running Time (milli-seconds)	76.65	89.15

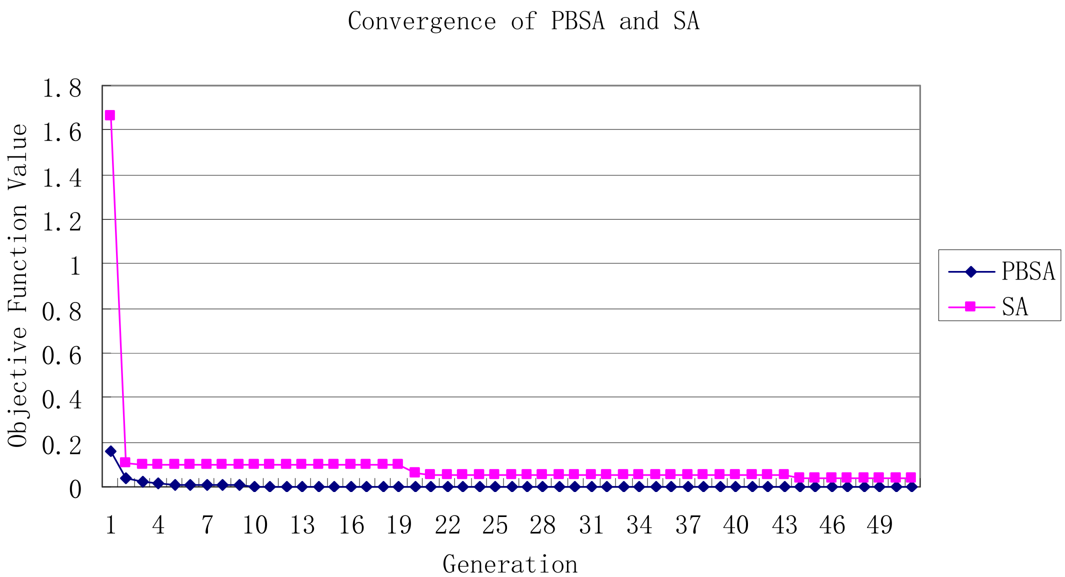


Figure 6.4 Comparison of Convergence (Mean of 20 Runs for Each Genration)

Table 6.2 Optimized Results with Different N/Kmax (20 runs)

$N=K_{max}$	PBSA		SA	
	Mean	Std. dv	Mean	Std. dv
20	0.0118504	0.0152637	0.0369192	0.0247537
40	0.0038489	0.0039957	0.0205690	0.0178339
60	0.0029787	0.0041941	0.0221259	0.0207025
80	0.0014236	0.0009594	0.0488890	0.0225573
100	0.0016211	0.0019797	0.0235211	0.0224531
150	0.0002274	0.0002740	0.0201639	0.0203240
200	0.0003077	0.0004287	0.0361220	0.0260059

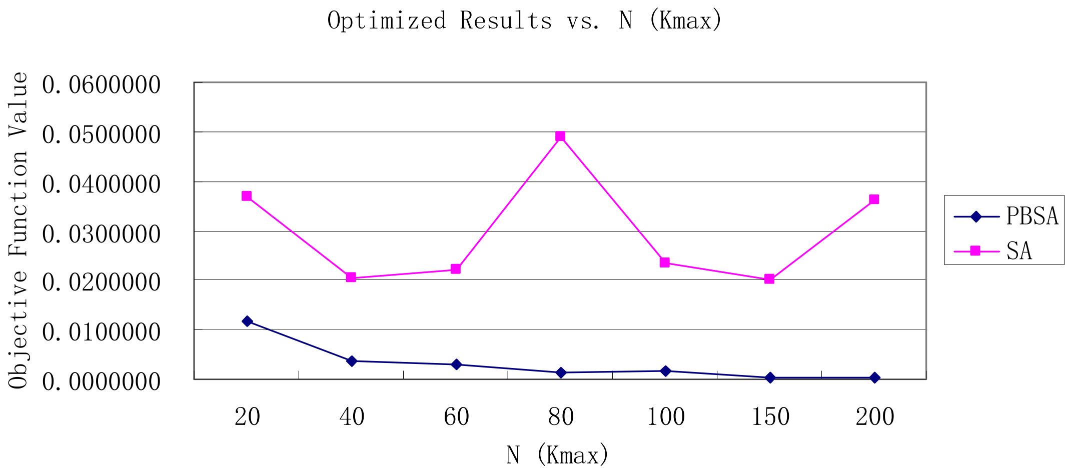


Figure 6.5 Optimized Results with Different N/Kmax (Mean of 20 runs)

6.4 Summary

In this chapter, a two-stage modified simulated annealing algorithm is developed. In the first stage, an initial optimization process is performed through a population-base simulated annealing (PBSA) approach and the result will be sent to the second stage as a relatively good initial solution. In the second stage, traditional simulated annealing (SA) method is used to do refined search inside the prominent region provided by the first step.

A multiple-optima function is used to compare the PBSA and the SA. It shows that PBSA dominates SA in the solution quality, run-to-run variance and running time, and that PBSA is sensitive to the population size.

Chapter 7 Work Zone Optimization based on Analytic Model for Time-Dependent Traffic Inflows

According to the previous developed work zone cost models (Chapter 3), work zone cost is quite sensitive to work duration, which relates to work zone length, work zone start time and the traffic distribution under time-dependent traffic inflows. Multiple local optima may occur in the solution space. Therefore, we cannot use differential calculus in locating the optimum points as we do for steady traffic inflows. A different methodology is needed to optimize the work zone cost for time-dependent traffic inflows.

7.1 Problem Statement

For time-dependent traffic inflows, efficient scheduling and traffic control strategy may significantly reduce the total cost, including agency cost and use cost. Based on time-dependent inflows, the issues considered in this chapter include:

(1) What is the best starting time and duration for a work zone within specific time constraints?

(2) What kind of lane closure strategy is preferable, depending on circumstances?

(3) Should the traffic be diverted if there are one or more detour routes available? What should be the diversion fractions be?

(4) What is the most cost-effective work rate and the corresponding work cost?

The statement of this work zone optimization problem for time-dependent traffic inflows can be formulated as:

The objective function:

$$J = \min C_t(\overrightarrow{X}) \tag{7-1}$$

The decision variables:

$$\overrightarrow{X} = \{Ts, Dw, \overrightarrow{Alt}, \overrightarrow{Rate}, \overrightarrow{P}\} \tag{7-2}$$

Subjective to following constraints:

(1) Time window constraint: $T_{s,\min} \le T_s \le T_{s,\min} + D_{w,\max}$, $T_{s,\min} \le T_s + D_w \le T_{s,\min} + D_{w,\max}$

(2) Work zone length constraint: $L_{w,\min} \le L_w \le L_{w,\max}$

(3) User specified lane closure alternatives: $\overline{Alt} \in \{\overline{Alt_1}, \overline{Alt_2}, ..., \overline{Alt_m}\}$

(4) User specified work zone rate parameters: $\overrightarrow{Rate} \in \{\overrightarrow{Rate}_1, \overrightarrow{Rate}_2, ..., \overrightarrow{Rate}_n\}$

(5) Maximum diverted fraction constrain: $\sum_i P_i \le P_{\max}$

where,

C_t = the total cost per completed lane mile maintenance work

T_s = work zone start time

D_w = work duration

$\overrightarrow{Alt}$ = lane closure alternative, which includes the information about the number of closed lane (n_{c1}), the number of usable counter flow lanes (n_{c2}) and the number of access lanes (n_{a1}).

$\overrightarrow{Rate}$ = work zone rate, which includes the information about the fixed setup time (z_1), average maintenance time per lane length (z_2), fixed setup cost (z_3) and average maintenance cost per lane length (z_4).

$\overrightarrow{P}$ = diverted fractions, which include the diverted fraction of the traffic flow along work zone link (p_1) and the diverted fraction of the counter flow (p_2).

$T_{s,\min}$ = the allowable earliest work zone start time

$D_{w,\max}$ = the allowable longest work duration

$L_{w,\min}$ = the allowable minimum work zone length

$L_{w,\max}$ = the allowable maximum work zone length

$P_{\max}$ = the allowable maximum diverted fraction

$\overrightarrow{Alt}_i$ = the i th lane closure alternatives. Users should input available alternatives discretely. $\overrightarrow{Alt}_i = \{n_{c1,i}, n_{c2,i}, n_{a1,i}\}$

$\overrightarrow{Rate}_i$ = the i th lane closure alternatives. Users should input available alternatives discretely. $\overrightarrow{Rate}_i = \{z_{1,i}, z_{2,i}, z_{3,i}, z_{4,i}\}$

The goal for this problem is to search for the optimal decision variable combinations $\overrightarrow{X}^* = \{T_s^*, D_w^*, \overrightarrow{Alt}^*, \overrightarrow{Rate}^*, \overrightarrow{P}^*\}$ which yield the global minimum total cost per lane length for the work zone project C_t^*.

The constraints are set from some operation consideration. $T_{s,\min}$ and $D_{w,\max}$ establish a time period in which the construction activity is allowed, shown in Figure 7.1. The work zone length (L_w), which can be derived from the work duration through the relation $D_w = z_3 + z_4 L_w$, should be longer than the minimum length required for the work space ($L_{w,\min}$) and shorter than the project length ($L_{w,\max}$). Considering the roadway capability of the detour, the maximum allowable fraction of traffic volumes along the work zone which can be diverted to the detour ($P_{\max}$) is also introduced.

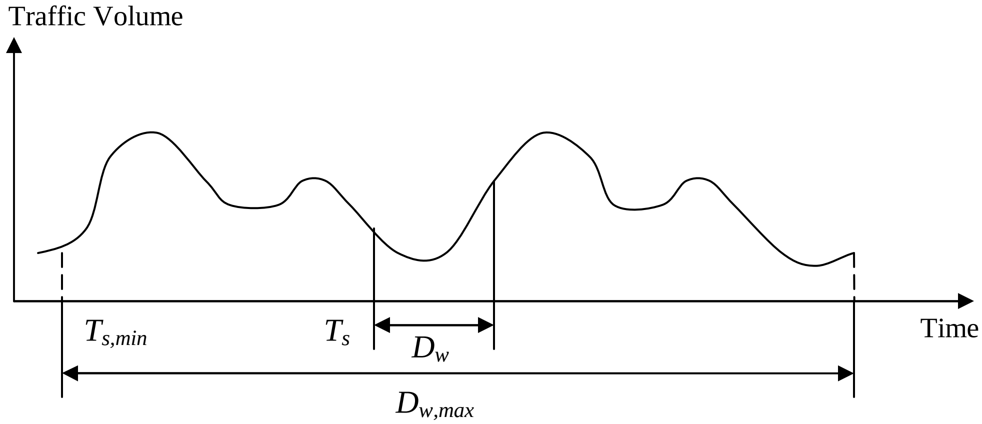


Figure 7.1 Permitted Work Time Period Defined by $T_{s,\min}$ and $D_{w,\max}$

7.2 Optimization Method

The general procedure to solve an optimization problem is shown in Figure 7.2. In this study, the two-stage modified simulated annealing algorithm proposed in Chapter 6 is applied as the solution search approach. In this section, we will present the other four components: (1) initial solution generation; (2) new solution generation; (3) solution evaluation and the inputs required in solution evaluation.

(1) Initial Solution Generation

A candidate solution vector $\vec{X}$ includes five components $\{Ts, Dw, \overrightarrow{Alt}, \overrightarrow{Rate}, \vec{P}\}$.

The lane closure alternative $\overrightarrow{Alt}$ and the work rate parameters $\overrightarrow{Rate}$ are randomly chosen from user specified data set $\{\overrightarrow{Alt_1}, \overrightarrow{Alt_2}, ..., \overrightarrow{Alt_m}\}$ and $\{\overrightarrow{Rate_1}, \overrightarrow{Rate_2}, ..., \overrightarrow{Rate_n}\}$. Considering that maintaining multiple lanes simultaneously may increase the work efficiency and therefore reduce cost as well as work time, work rate parameters $\overrightarrow{Rate} = \{z_1, z_2, z_3, z_4\}$ can be adjusted according to the selected $\overrightarrow{Alt}$ and $\overrightarrow{Rate}$ by employing a cost deduction factor (F_c) and a time deduction factor (F_t). These two adjust factors should be provided by users.

Based on hourly traffic distribution along the work zone link, T_s and D_w are always generated to make work period between two traffic peak hours. A check procedure is added to adjust the generated T_s and D_w so that: (1) the derived L_w can satisfy the work zone length constraints; (2) the derived L_w can provide an integer number of zones, with which the project length (L_{max}) can be divided as evenly as possible, and (3) T_s and D_w can satisfy the work zone time constraints.

The diverted fraction vector $\overline{P}$ is randomly generated between zero and P_{max}. If there is more than one diverted fraction, a check procedure is added to insure not only that each diverted fraction can satisfy the constraint but also that the sum of all diverted fractions is less than P_{max}.

(2) New Solution Generation

New solutions are generated in three ways.

(a) With a probability, a neighborhood solution is generated from the old solution through modifying one or more decision variables included in the old solution.

(b) With a probability, a new solution is generated in the same way as the initial solutions, without considering the information in the old solution.

(c) With a probability, a new solution is generated from the current best solution ever found through modifying one or more decision variables included in the best solution.

The same solution check procedure as that in initial solution generation part is applied to make the new solution satisfy all constraints.

(3) Solution Evaluation

Each solution is evaluated by calculating the total cost per completed lane length work with the following expression:

$$C_t = \frac{C_T}{L_w \cdot n_w} = \frac{C_M + C_D + C_A}{L_w \cdot n_w} = \frac{z_1 + z_2 L_w n_w + (v_d + \frac{v_a n_a}{10^8}) t_d}{L_w \cdot n_w} \quad (7\text{-}3)$$

where,

n_w = the number of maintained lanes

L_w = the work zone length;

C_M = the maintenance cost for a work zone

C_D = the user delay cost for a work zone

C_A = the accident cost for a work zone

n_a = represents the number of accidents per 100 million vehicle hour

v_a = the average cost per accident

v_d = the average user's time value

t_d = the user delay

In this chapter, we use the analytic method for time-dependent traffic inflows, which is developed in Section 3.3 in Chapter 3, to calculate the user delay t_d.

Besides the parameters which have been mentioned above, the following inputs are required for solution evaluation:

(a) Geometrics of road network.

(b) Traffic data including hourly traffic flow volumes along mainline route and detour route and the parameters describing speed, capacity and density.

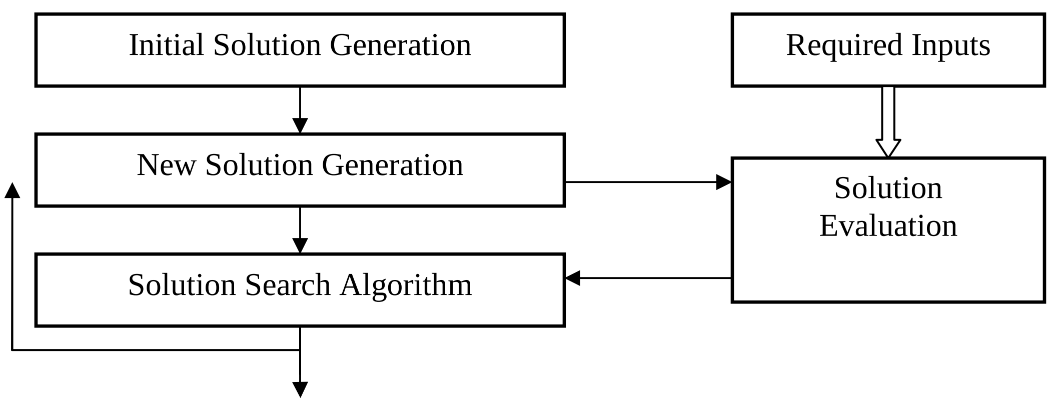


Figure 7.2 General Framework of an Optimization Procedure

7.3 Numerical Examples

In this section, a numerical experiment is performed to test our proposed optimization method for time-dependent traffic inflows. The effects of various parameters on the optimized results are examined.

7.3.1 Network Description

A work zone on a two-lane two-way highway considering single detour is analyzed (Figure 7.3).The baseline numerical values are shown in Table 7.1. The baseline values are data or estimates provided by the Maryland State Highway Administration. Table 7.2 and 7.3 show the available lane closure alternatives and work rate parameters. Table 7.4 shows the assumed traffic distributions on the mainline and detour over each day.

Two time constraints are tested: (1) daytime window, 5:00 – 20:00, which is defined by $T_{s,\min}$ =5 and $D_{w,\max}$ =15 hours; (2) two-day time window, 0:00 the first day-24:00 the next day, which is defined by $T_{s,\min}$ =0 and $D_{w,\max}$ =48 hours. The other constraints are: (1) $L_{w,\min}$ =0.1 mile; (2) $L_{w,\max}$ =2.5 mile; (3) $P_{\max}$ =30%.

Nine Traffic levels with traffic volume multipliers ranging from 0.2 to 1.8 for mainline traffic volumes are tested. For each traffic level, the distribution of the traffic flow keeps unchanged while the hourly volume increase or decreases to the production of the baseline volume and the traffic level multiplier.

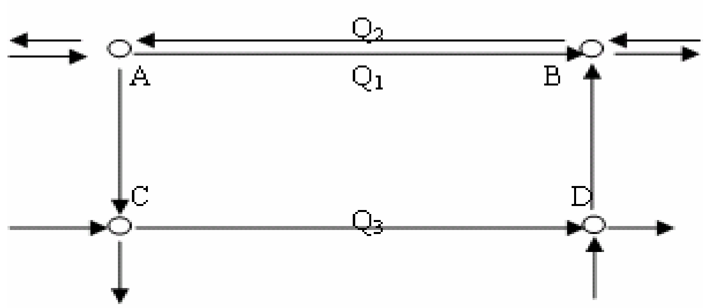


Figure 7.3 The Roadway Network and Its Simplified Model

Table 7.1 Notation and Baseline Numerical Inputs

Variable	Description	Value
L_{AB}	Length of Segment AB	3.11 miles
L_{AC}	Length of Segment AC	2.49 miles
L_{CD}	Length of Segment CD	0.93 miles
L_{DB}	Length of Segment DB	0.93 miles
NAB	Number of lanes in Segment AB	2 lanes
NCD	Number of lanes in Segment CD	1 lane
Kj	Jam Density	200 veh/mile
c_0	Maximum discharge rate without work zone	2,200 vph /lane
c_w	Maximum discharge rate with work zone	1,600 vph /lane
V_{AB}	Average approaching speed	65 mph
Vw	Average work zone speed	45 mph
V_{CD}	Free flow speed in Segment CD	45 mph
$V_{AC/DB}$	Average speed in Segment AC/DB	45 mph
T_{int}	Average waiting time passing intersections along the detour	30 seconds/veh
n_a	Number of crashes per 100 million vehicle hours	40 acc/100mvh
v_a	Average accident cost	142,000$/accident
v_d	Value of user time	12 $/veh·hr
z_1	Fixed setup cost	$/zone
z_2	Average maintenance cost per lane kilometer	$/lane.mile
z_3	Fixed setup time	hr/zone
z_4	Average maintenance time per lane kilometer	hr/lane.mile
n_{c1}	Number of closed lanes in Direction 1	lanes
n_{c2}	Number of usable counter flow lanes in Direction 2	lanes
n_{a1}	Number of access lanes	lanes
n_w	Number of maintained lanes	lanes

Table 7.2 The Available Lane Closure Alternatives

Alternative		n_{c1}	n_{c2}	n_{a1}	n_w
1	One Lane Closure	1	0	0	1
2	Two Lane Closure with Crossover	2	1	0	2

Table 7.3 The Available Work Rates

Work Rate	z_1	z_2	z_3	z_4
Slow	1000	32000	2	12
Normal	1000	33000	2	10
Fast	1000	34000	2	8

Table 7.4 AADT and Hourly Traffic Distribution on a Two-Lane Two-Way Freeway

Time Period	Time	Mainline		Detour
		Q1	Q2	Q3
0	0:00-1:00	220	930	392
1	1:00-2:00	157	645	391
2	2:00-3:00	148	301	367
3	3:00-4:00	198	238	432
4	4:00-5:00	448	240	432
5	5:00-6:00	1425	326	432
6	6:00-7:00	2941	580	734
7	7:00-8:00	3541	887	1276
8	8:00-9:00	2897	977	1505
9	9:00-10:00	2509	1134	1363
10	10:00-11:00	1793	1283	951
11	11:00-12:00	1586	1589	772
12	12:00-13:00	1528	1544	700
13	13:00-14:00	1475	1673	670
14	14:00-15:00	1541	2074	773
15	15:00-16:00	1414	2808	954
16	16:00-17:00	1079	3501	1042
17	17:00-18:00	957	3719	1026
18	18:00-19:00	991	3061	832
19	19:00-20:00	779	2171	770
20	20:00-21:00	554	1433	644
21	21:00-22:00	504	1314	559
22	22:00-23:00	436	905	392
23	23:00-24:00	325	720	391
AADT		114016	113586	19528
Average Hourly Volume		4751	4733	813

Table 7.5 Traffic Levels Ranging from 1.2 to 1.8

Traffic Level	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8
Q1 AADT	5889	11778	17667	23556	29446	35335	41224	47113	53002
Q2 AADT	6810	13621	20431	27242	34053	40863	47674	54484	61295

7.3.2 Convergence Analysis

Figure 7.4 shows the convergence of the optimization process for two-day time constraint and daytime constraint under baseline traffic volumes. Since the algorithm records the best solution ever found at each generation, a monotonically decreasing relation appears in the figure. We see that each optimization process converges to an optimized solution. The optimized work zone cost within daytime constraint is higher than the optimized work zone cost within two-day time constraint. This is consistent with our expectation because daytime window is a more restrictive constraint than the two-day time window.

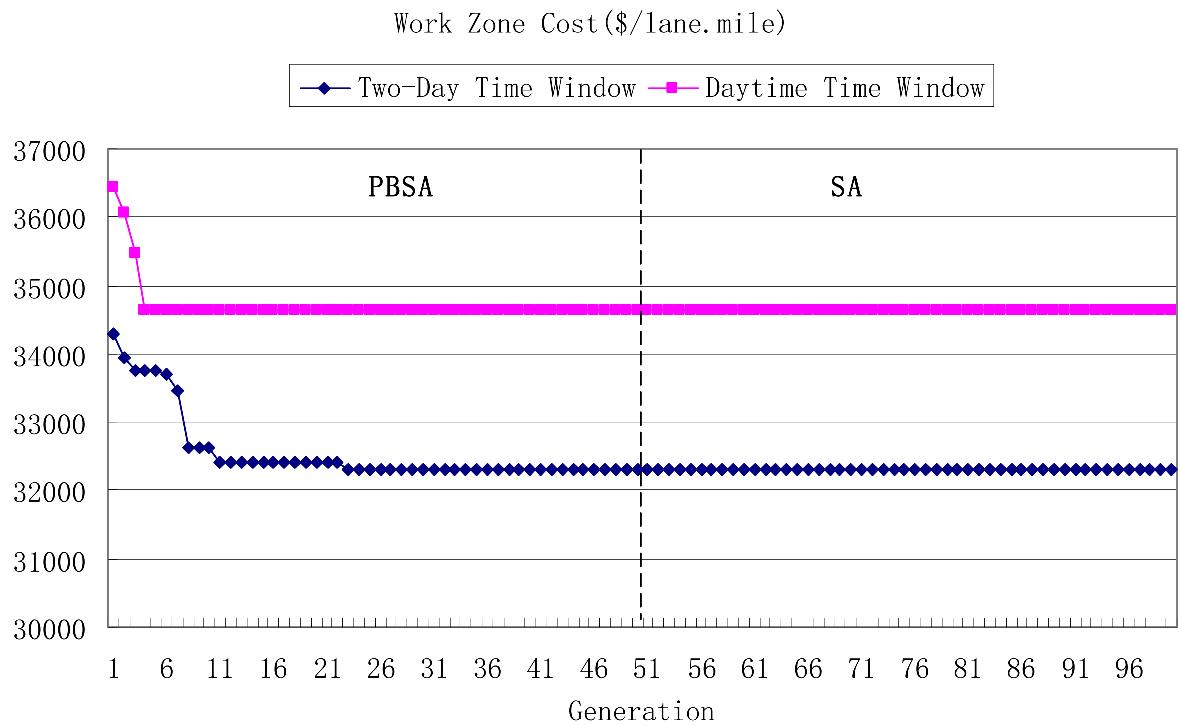


Figure 7.4 Convergences for Baseline Volume

7.3.3 Reliability Analysis

A hundred replications are performed to optimize the work zone plan for the baseline traffic volume. The optimized results are shown in Figure 7.5. The average running time is 6.25 seconds. The statistics of the optimized work zone costs are shown in Table 7.6. A small coefficient of variance indicates that this optimization algorithm is reliable.

Table 7.6 Statistics of the Optimized Work Zone Cost ($/lane.mile)

Rep.	Mean	Max	Min	Std. Dv.	Coef. of Variance
100	32565	33676	32296	535	1.6%

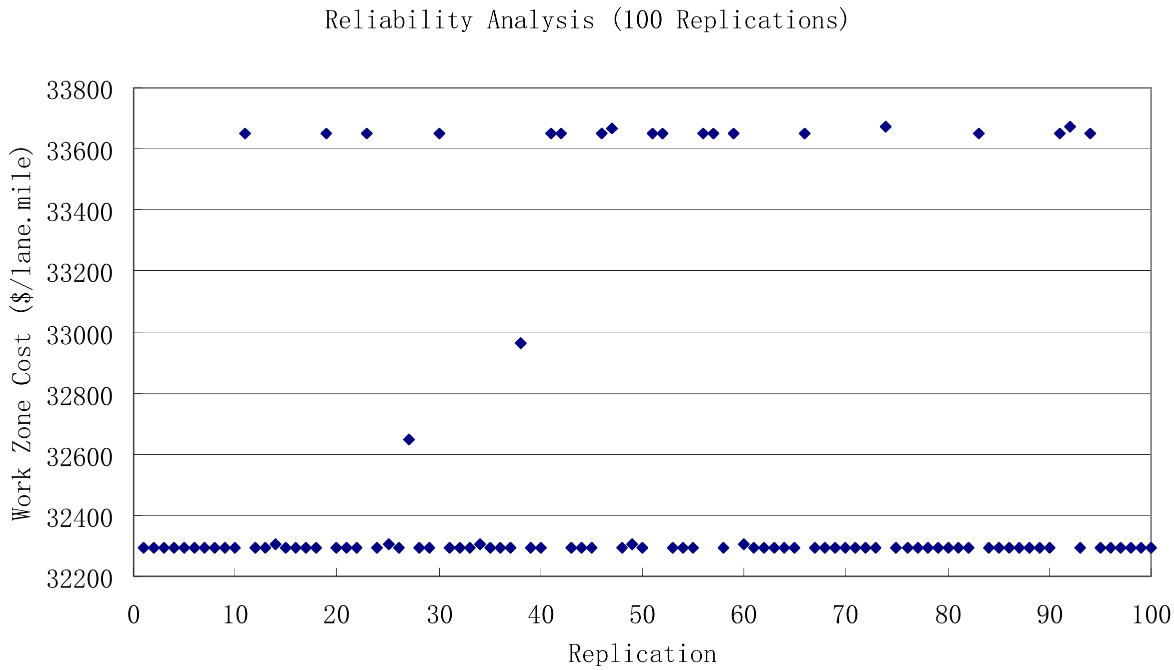


Figure 7.5 Reliability of the Optimization Algorithm

7.3.4 Sensitivity Analysis

In this case, we seek to examine the impact of traffic volumes and time constraints on the optimized results. Optimization is performed for scenarios with different traffic level and time windows. The same algorithm parameters are used in the two-stage modified simulated annealing algorithm. In the first stage, the initial temperature is 5000, the stop temperature is 0, the step size of temperature size is 100 and the population size is 200. In the second stage, the initial temperature is 500, the stop temperature is 0, the step size of temperature size is 10 and the repetition factor is 5. Each scenario is re-optimized in ten replications. The best solutions are recorded in Tables 7.6 and 7.7.

Figure 7.6 shows all the optimized work zone costs. As expected, the optimized work zone cost within daytime constraint is higher than that with two-day time constraint at all traffic levels.

(1) Two-Day Time Window ($T_{s,\min}$ =0 and $D_{w,\max}$ =48 hours.)

The optimized solutions in traffic level ranging from 0.2 to 1.8 are displayed in Table 7.7. As the traffic increases, the impact of the work zone activity on motorists also increases. At traffic levels below 1.8, the optimized lane closure alternative has two closed lanes and a lane crossover to the counter-flow lane. Traffic flows in both directions (Q_1 and Q_2). As traffic increases, the optimized work zone cost per lane mile increases, the optimal work zone starting time moves to low-volume late night hours, and the work duration as well as work zone length decrease. Note that at traffic level 1.8, the two lane closure is no longer affordable and the ten hour one lane closure becomes the best solution.

Table 7.7 Optimized Solutions within a Two-Day Time Window

Traffic Level	Ts	Dw (hr)	Alt ID	Work Rate	P (%)	Lw (mile)	Cm ($/ln.ml)	Cd ($/ln.ml)	Ca ($/ln.ml)	Ct ($/ln.ml)
0.2	0:00	30	2	Slow	0	1.22	30809	562	0.26	31371
0.4	19:00	17	2	Slow	0	0.66	31160	592	0.28	31753
0.6	19:00	11	2	Slow	0	0.39	31667	355	0.17	32021
0.8	20:00	10	2	Slow	0	0.35	31825	376	0.18	32201
1.0	20:00	10	2	Slow	0	0.35	31825	470	0.22	32296
1.2	21:00	8	2	Slow	0	0.26	32300	380	0.18	32680
1.4	22:00	7	2	Slow	0	0.22	32680	339	0.16	33019
1.6	22:00	7	2	Slow	0	0.22	32680	387	0.18	33068
1.8	19:00	10	1	Slow	0	0.67	33500	556	0.26	34057

(2) Daytime Time Window ($T_{s,\min}$ =5 and $D_{w,\max}$ =15 hours.)

Table 7.8 shows the optimized solutions in increasing traffic levels within a daytime 15-hour time window,

In this time window, the work zone activity is limited to a daytime period from 5:00 am to 20:00 pm in the same day. The optimized work zone cost per lane mile is more sensitive to traffic level in the day time window. When the traffic level increases, the work zone duration as well as work zone length decreases. At high traffic levels, a detour and a faster work rate are desirable.

Table 7.8 Optimized Solutions within a Daytime 15-hour Time Window

Traffic Level	Ts	Dw (hr)	Alt ID	Work Rate	P (%)	Lw (mile)	Cm ($/ln.ml)	Cd ($/ln.ml)	Ca ($/ln.ml)	Ct ($/ln.ml)
0.2	5	15	2	slow	0	0.57	31277	441	0.21	31717
0.4	5	15	2	slow	0	0.57	31277	882	0.42	32159
0.6	9	11	1	slow	0	0.75	33333	770	0.36	34104
0.8	10	10	1	normal	0	0.67	33500	862	0.41	34363
1.0	11	9	1	fast	0	0.58	33714	931	0.44	34646
1.2	16	4	1	fast	0	0.25	38000	374	0.18	38374
1.4	16	4	1	fast	10	0.25	38000	437	0.21	38437
1.6	16	4	1	fast	10	0.25	38000	2685	1.27	40687
1.8	16	4	1	fast	20	0.25	38000	5701	2.70	43704

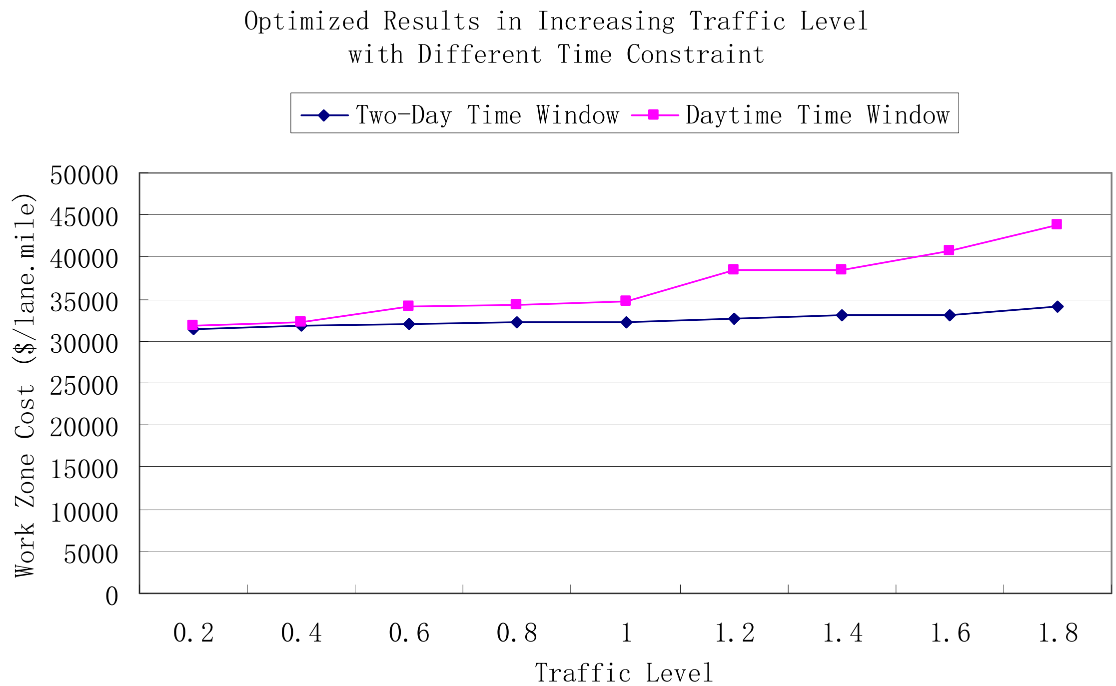


Figure 7.6 Optimized Work Zone Cost vs. Increasing Traffic Levels

7.4 Summary

In this chapter, a methodology to solve the work zone optimization problem based on the analytic model for time-dependent traffic inflows is proposed and tested through a numerical example. The optimization model applies the two-stage modified SA algorithms, which is developed in Chapter 6 as the optimization algorithm. The procedures for solution generation, modification and evaluation are also presented.

Chapter 8 Work Zone Optimization based on Simulation

In Chapter 6, an optimization method, which uses the two-stage modified SA algorithm to search for the optimal combination of decision variables in the solution space, is developed to solve the work zone optimization problem.

With the same methodology, simulation method, instead of the analytic method for time-dependent traffic inflows, is applied to evaluate the objective function in the optimization process. The work zone optimization model based on simulation is presented in this chapter.

However, such optimization through simulation may impose severe computation burdens. To make the search algorithm as efficient as possible and thus reduce the computational efforts to a more acceptable level, a hybrid approach, which combines simulation and the analytic method, is proposed in this chapter.

8.1 Problem Statement

The statement of the work zone optimization based on simulation model is the same as the problem defined in Chapter 6:

The objective function:

$$J = \min C_t(\overrightarrow{X}) \tag{8-1}$$

The decision variables:

$$\overrightarrow{X} = \{Ts, Dw, \overrightarrow{Alt}, \overrightarrow{Rate}, \overrightarrow{P}\} \tag{8-2}$$

The constraints:

(1) Time window constraint: $T_{s,\min} \le T_s \le T_{s,\min} + D_{w,\max}$, $T_{s,\min} \le T_s + D_w \le T_{s,\min} + D_{w,\max}$

(2) Work zone length constraint: $L_{w,\min} \le L_w \le L_{w,\max}$

(3) User specified lane closure alternatives: $\overrightarrow{Alt} \in \{\overrightarrow{Alt_1}, \overrightarrow{Alt_2}, ..., \overrightarrow{Alt_m}\}$

(4) User specified work zone rate parameters: $\overrightarrow{Rate} \in \{\overrightarrow{Rate_1}, \overrightarrow{Rate_2}, ..., \overrightarrow{Rate_n}\}$

(5) Maximum diverted fraction constrain: $\sum_i P_i \le P_{\max}$

where C_t = the total cost per completed lane mile

T_s = work zone start time

D_w = work duration

$\overrightarrow{Alt}$ = lane closure alternative, which includes the information about the number of closed lanes (n_{c1}), the number of usable counter-flow lanes (n_{c2}) and the number of access lanes (n_{a1}).

$\overrightarrow{Rate}$ = work zone rate, which includes the information about the fixed setup time (z_1), average maintenance time per lane length (z_2), fixed setup cost (z_3) and average maintenance cost per lane length (z_4).

$\overrightarrow{P}$ = diverted fractions, which include the diverted fraction of the traffic flow along work zone link (p_1) and the diverted fraction of the counter flow (p_2).

$T_{s,\min}$ = the earliest allowable work zone start time

$D_{w,\max}$ = the longest allowable work duration

$L_{w,\min}$ = the minimum allowable work zone length

$L_{w,\max}$ = the maximum allowable work zone length

$P_{\max}$ = the maximum allowable diverted fraction

$\overrightarrow{Alt}_i$ = the i th lane closure alternatives. User should input available alternatives discretely. $\overrightarrow{Alt}_i = \{n_{c1,i}, n_{c2,i}, n_{a1,i}\}$

$\overrightarrow{Rate}_i$ = the i th lane closure alternatives. User should input available alternatives discretely. $\overrightarrow{Rate}_i = \{z_{1,i}, z_{2,i}, z_{3,i}, z_{4,i}\}$

8.2 Optimization Method

The optimization method includes four major elements: (1) Initial solution generation; (2) New solution generation; (3) Solution evaluation and (4) Solution search algorithm. All parts except for the third element, solution evaluation, in the work zone optimization model based on simulation are the same as those in the optimization model based on analytic method.

8.2.1 Solution Evaluation Based on Simulation Model

Each solution is evaluated by calculating the total cost per completed lane length with the following expression:

$$C_t = \frac{C_T}{L_w \cdot n_w} = \frac{C_M + C_D + C_A}{L_w \cdot n_w} = \frac{z_1 + z_2 L_w n_w + (v_d + \frac{v_a n_a}{10^8}) t_d}{L_w \cdot n_w} \quad (8\text{-}3)$$

where,

n_w = the number of maintained lanes

L_w = the work zone length;

C_M = the maintenance cost for a work zone

C_D = the user delay cost for a work zone

C_A = the accident cost for a work zone

n_a = the number of accidents per 100 million vehicle hours

v_a = the average cost per accident

v_d = the average user's time value

t_d = the user delay

In this study, simulation based on CORSIM is performed to evaluate the user delays (t_d) caused by the work zone activity, whose characteristics are specified in a candidate solution. The procedure for simulating different work zone characteristics has been introduced in Chapter 4. However, to generate simulation input files and obtain simulation output in an automated way, three modules are needed to link the optimization process with CORSIM model, shown in Figure 8.1. The three modules are the preparation module, the preprocessor module and the postprocessor module.

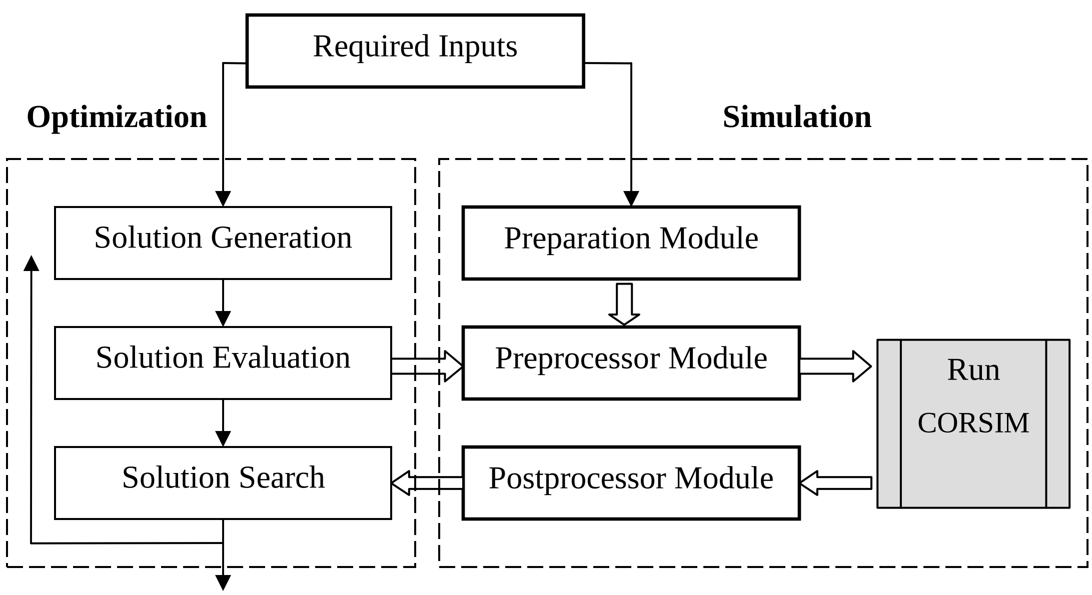


Figure 8.1 Links between Optimization Process and Simulation Process

1. **Required Inputs**

Before evaluating solutions, users must provide the following inputs:

(1) The value of all parameters included in Eq. (8-3);

(2) Two CORSIM input files with the format of *.trf file, which provide datasets describing geometrics of the study network, 24-hour traffic information and traffic control parameters. Each input file includes 12 time periods with 1 hour for each time period. Hourly time-varying traffic information from 0:00 to 12:00 is recorded in the first CORSIM input file (Morning 12-hour Simulation Input File). Hourly time-varying traffic information from 12:00 to 24:00 is recorded in the second CORSIM input file (Afternoon12-hour Simulation Input File).

2. **Preparation Module**

The Preparation Module is used to provide some of required data needed in the preprocessor and postprocessor modules. The framework of the preparation module is displayed in Figure 8.2.

Step1: For the Morning 12-hour Simulation Input File and the Afternoon 12-hour Simulation Input File, call CORSIM.DLL to run simulation. Two output files can be obtained after the simulation is completed.

Step 2: From the output files, get hourly traffic volumes in each link and hourly network-widely delay time. The former will be used to get peak hours in the solution generation process and to calculate new turn movement percentages with detours in the preprocessor module. The latter will be used in the postprocessor module to calculate user delays in a normal situation without a work zone.

3. **Preprocessor Module**

The purpose of the Preprocessor Module is to generate new CORSIM input files according to the work zone information in the candidate solution generated from the optimization process. Figure 8.3 shows the flow chart of the preprocessor module.

Step 1: According to the work zone characteristics provided by the solution, calculate the total time period need to simulate.

Step 2: Due to the limitation of up to 19 time periods in CORSIM, more than one input file may have to be generated for simulating the work zone activity. Based on the Morning 12-hour Simulation Input File and the Afternoon 12-hour Simulation Input File, generate the input files with different simulation start time and period. Note that in these input files no work zone information is recorded.

Step 3: According to the work zone information in the solution, modify the input files generated in step 2. The details of the modification procedure have been introduced in Chapter 4. After the modifications, new CORSIM input files with work zone information can be obtained.

Step 4: For these new CORSIM input files, call CORSIM.DLL to run simulation.

4. **Postprocessor Module**

The objective of the Postprocessor Module is to interpret the CORSIM outputs to the objective function values, which should be send back to the optimization process. The steps are demonstrated in Figure 8.4.

Step 1: Read the network-widely delay times from the simulation outputs of the CORSIM input files generated in the preprocessor module.

Step 3: Calculate the user delay in work zone conditions.

Step 4: Calculated the user delay in a normal situation without work zones according to the and hourly network-widely delay time obtained in the preparation module.

Step 5: Calculate the used delay caused by work zone activity defined in the candidate solution by subtracting the delay without work zones from the delay with work zones.

Step 6: Calculate the work zone total cost per lane length with Eq. (8-3).

8.2.2 Optimization Based On Simulation Method

The two-stage modified simulated annealing algorithm developed in Chapter 6 is applied to solve the work zone optimization problem considering time-varying traffic conditions.

We can use the simulation method to evaluate the objective function in both the initial optimization stage and the refined optimization stage. Then the optimization process is wholly based on simulation.

However, as we know, microscopic simulation is quite time-consuming. Depending on the network size and traffic congestion level, the running time for one-hour simulation ranges from several seconds to several minuets. For example, assuming the average time to simulate one solution is 5 minutes, it will take 3.47 days to finish evaluating 1000 solutions. Besides, there is still a risk of obtaining a sub-optimal solution because CORSIM may underestimate delays for an over-saturated network.

8.2.3 Optimization through Hybrid Method

In order to reduce the computational burden while still estimating precise work zone costs, a hybrid method is proposed in this subsection. In this method, the analytic method for time-dependent traffic flows is applied to evaluate the objective function in the first stage of the two-stage modified SA algorithm. Initial optimization based on the analytic model using PBSA is performed and the result will be sent to the second stage as a relatively good initial solution. In the second stage, the optimization model based on simulation method uses SA to perform a refined search inside the promising region provided by the first stage.

Through this hybrid approach, complete simulations may be avoided in the early search phases. The optimizing search process based on simulation method can start from pre-optimized decision variables and thus may be able to reach a high quality optimized solution in an efficient way. The hybrid method will thus combine the benefits of the PBSA algorithm and the SA algorithm as well as integrate the advantages of macroscopic analytic methods and microscopic simulation methods.

Before optimizing through the hybrid method, users must provide traffic data along mainline and detour routes. The estimated parameters describing speed, capacity and density are also required to input.

To maintain the consistency between the analytic model and simulation models, .the study network should be simplified into the networks explored with the analytic models (such as the networks shown in Figure 3.3 in Chapter 3).

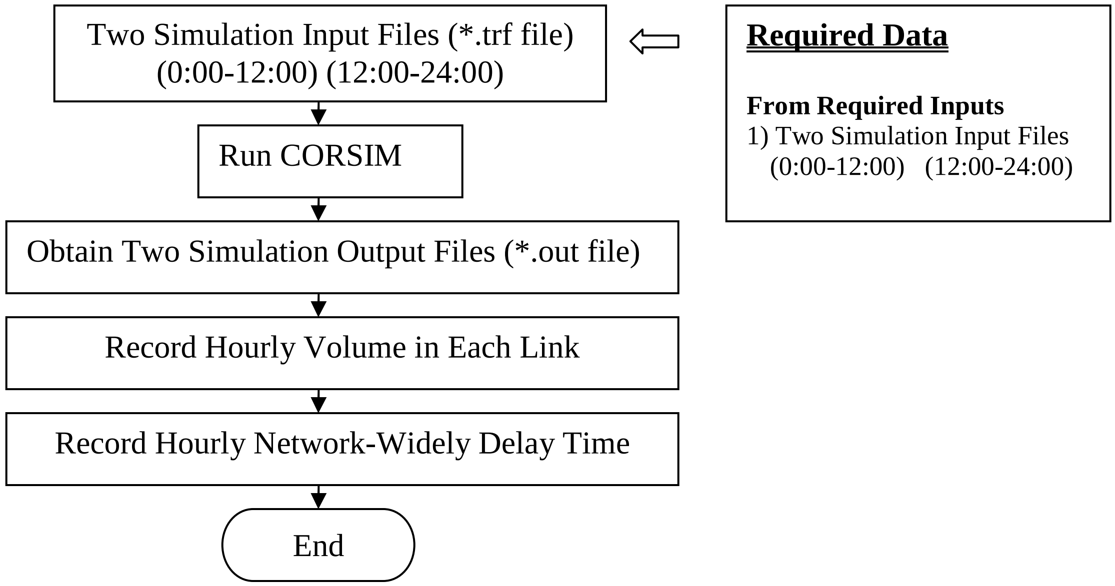


Figure 8.2 Framework of the Preparation Module

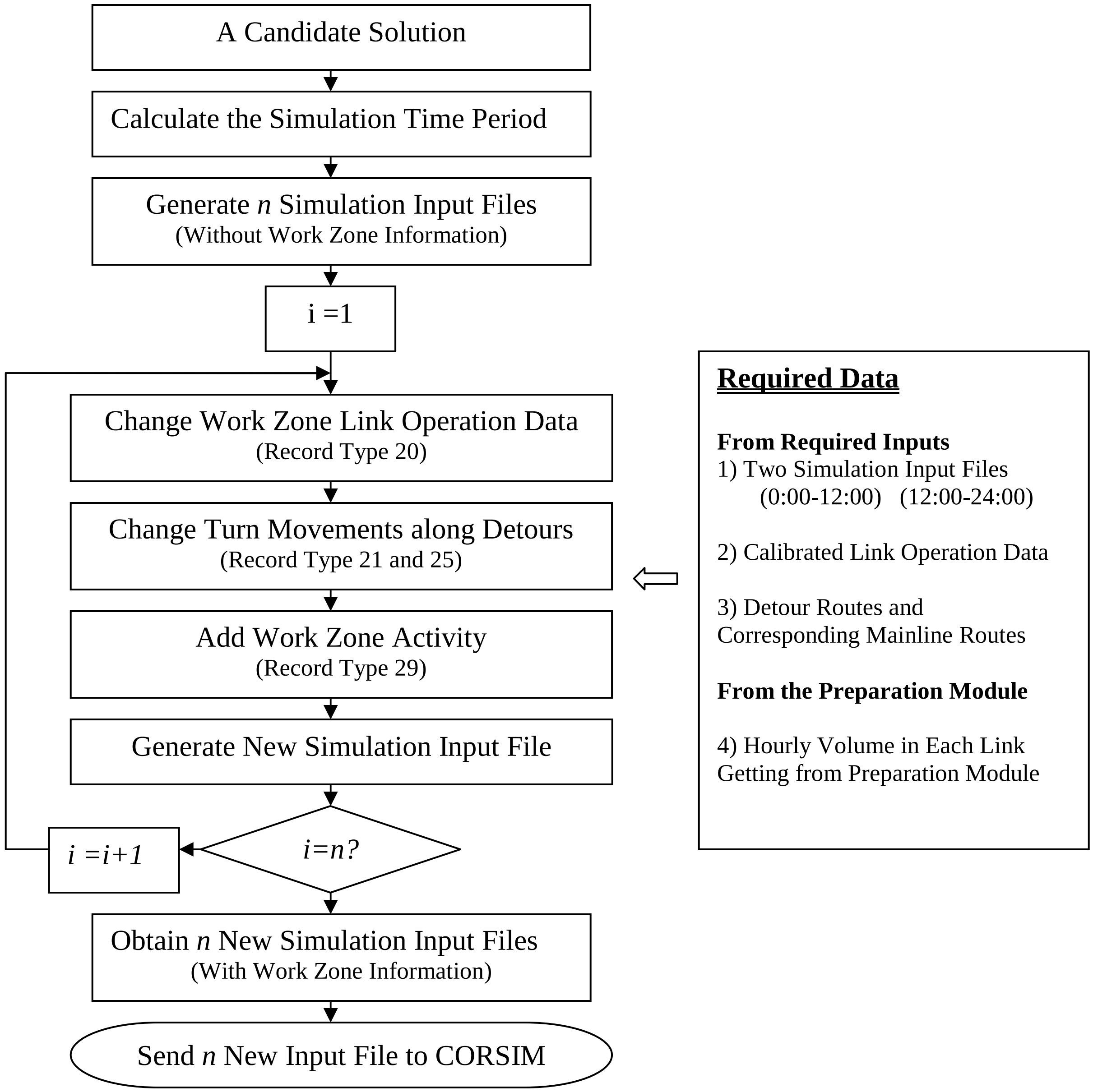


Figure 8.3 Framework of the Preprocessor Module

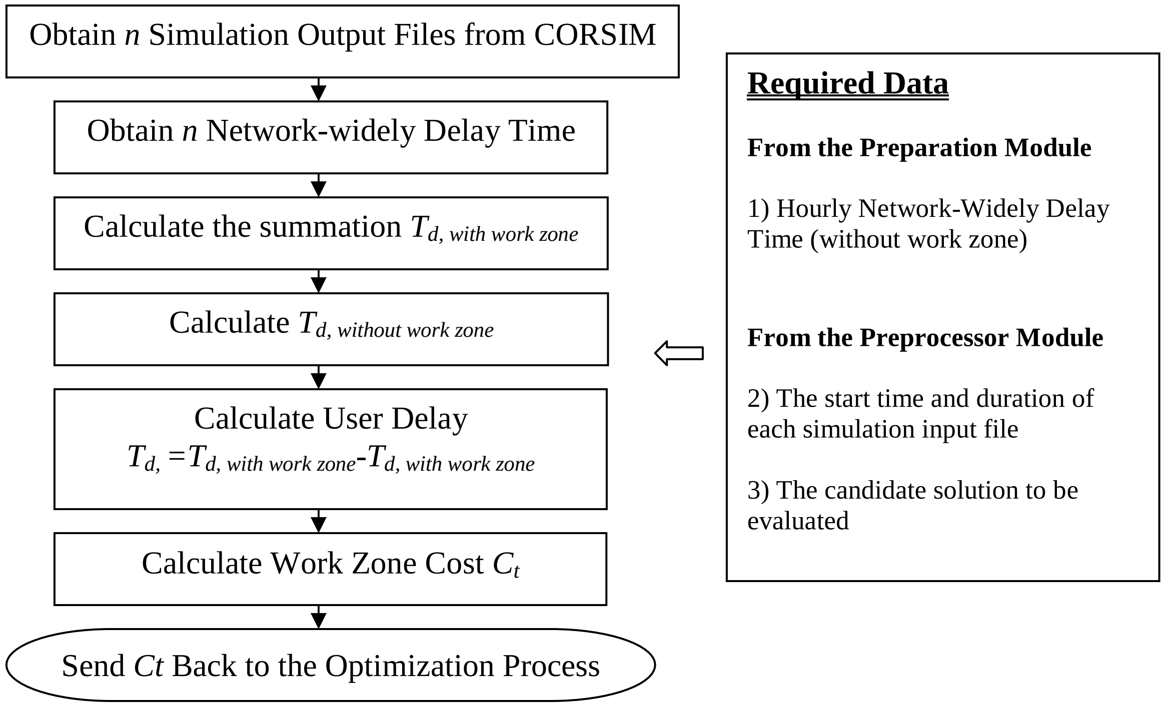


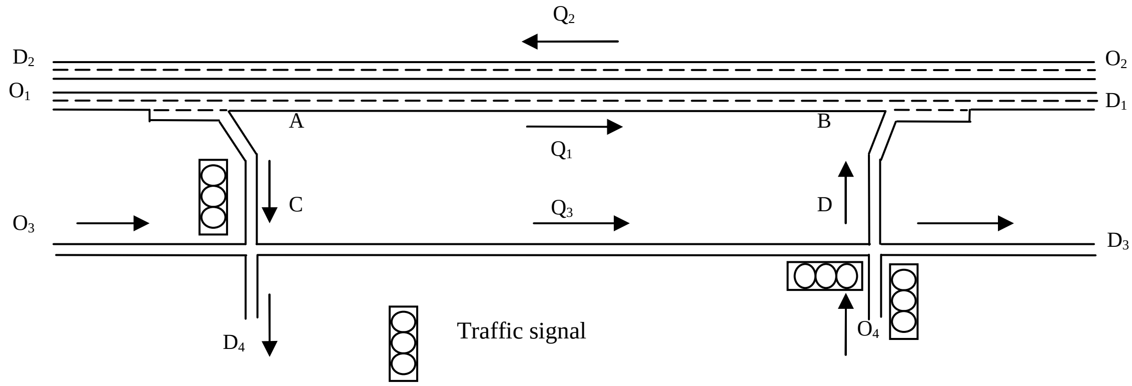
Figure 8.4 Framework of the Postprocessor Module

8.3 Numerical Examples

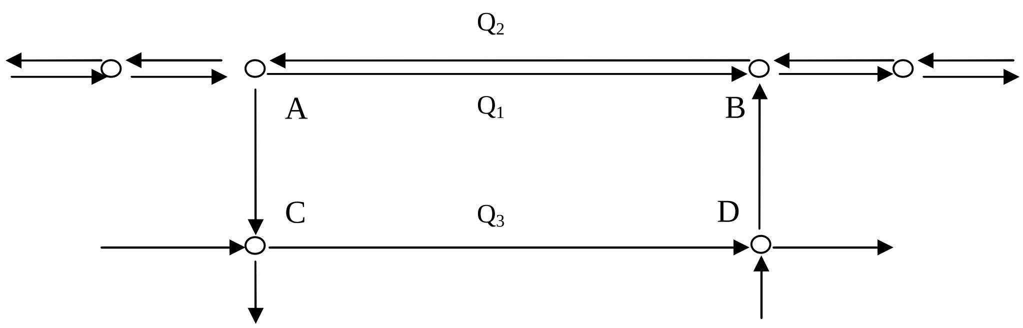
8.3.1 Network Description

A simple hypothetical network with multiple origins and destinations, shown in Figure 8.5 (a), is conceived in order to demonstrate the methodologies presented in this chapter. The network consists of a corridor with a four-lane two-way freeway and a parallel arterial. The freeway is 3.11 miles long. Both off-ramp deceleration lanes and on-ramp acceleration lanes are 800 feet long. The single-lane arterial is unidirectional. An actuated signal alternates permission between the off-ramp and the arterial. The arterial approaches to the on-ramp are controlled by a pre-timed signal control.

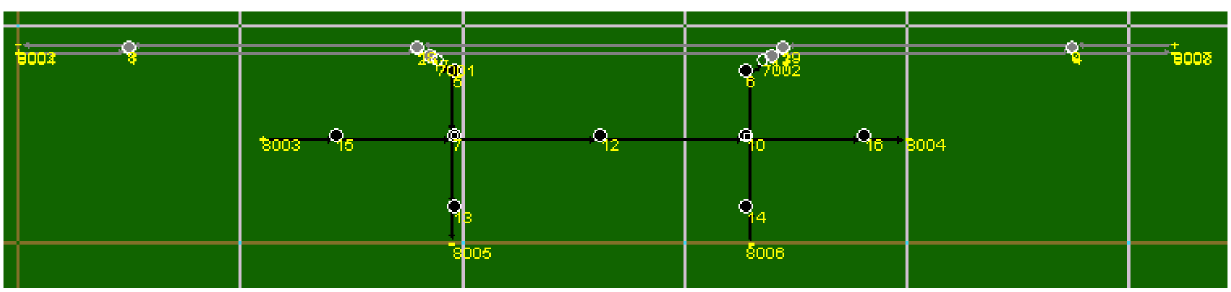
A one-lane maintenance project in the two-lane two-way highway segment AB is analyzed, assuming there is a detour route available (AC-CD-DB). An analytic model and a simulation model in CORSIM are established for the study road network, displayed in Figure 8.5 (b),(c).



(a) The Study Road Network



(b) The Analytic Model of the Study Road Network



(c) The Simulation Model of the Study Road Network in CORSIM

Figure 8.5 The Analytic Model and Simulation Model of the Study Network

The geometrics of the network and parts of variables related parameters are the same as in the numerical example tested in Chapter 7. More detailed representations of the study network including signal control plans and turning movements are provided in the simulation model in CORSIM.

We also use the same alternative work rate parameters and traffic volumes on the mainline and detour routes. Their baseline numerical values are shown in Tables 7.1, 7.3 and 7.4. Since one lane needs to be maintained, only one-lane closures are considered, thus assuming that no additional access lane is necessary.

The two-day time constraint ($T_{s,\min}$ =0 and $D_{w,\max}$ =48 hours) is tested in this example.

8.3.2 Optimization Results

Three work zone optimization models, based on simulation method, analytic method and hybrid method, are applied for the study network. They are run on a Pentium® 4 3.60 GHz PC with 0.99 GB of RAM. For the optimization based on simulation, the number of replications is one and the rubbernecking factor is set as 20% to obtain a work zone capacity of 1600 vph/lane.

Each optimization method is run three times and the best results are shown in Table 8.1. We can see the optimized solutions are quite close to each other. In fact, the difference between the work duration may come from the stochastic characteristic of simulation models. However, the running time for the simulation-based optimization is much greater than for the other two methods, even with fewer solution evaluations.

Table 8.1 Optimized Results from Three Optimization Models

Results	**Optimization based on Simulation Method**	**Optimization through Hybrid Method**	**Optimization based on Analytic Method**
Work Start Time	16:00	16:00	16:00
Work Duration	12 hours	12 hours	13 hours
Work Rate	Slow	Slow	Slow
Diverted Fraction	0%	0%	0%
Number of Solution Evaluation in PBSA	500	10000	10000
Number of Solution Evaluation SA	50	50	50
Total Generation	100	100	100
Running Time	114106.71 seconds	12461.36 seconds	5.48 seconds

The convergences of the three optimization processes are shown in Figures 8.6, 8.7 and 8.8. It is shown that in the first stage PBSA algorithm provides relatively good solutions in all three cases. In the hybrid optimization, the SA algorithm works quite efficiently to search for a better solution through simulation within the relatively good neighborhood provided by the first stage.

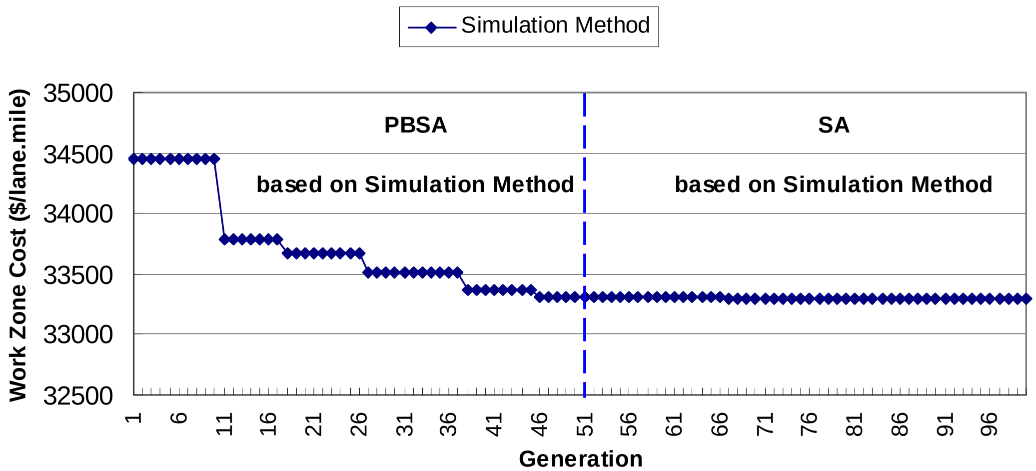


Figure 8.6 Optimization Convergence based on Simulation Method

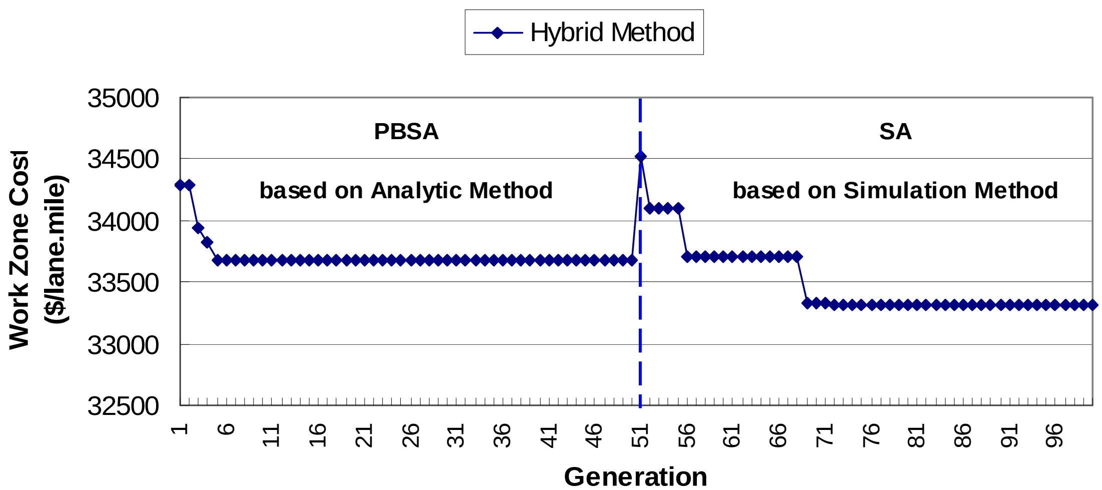


Figure 8.7 Optimization Convergence through Hybrid Method

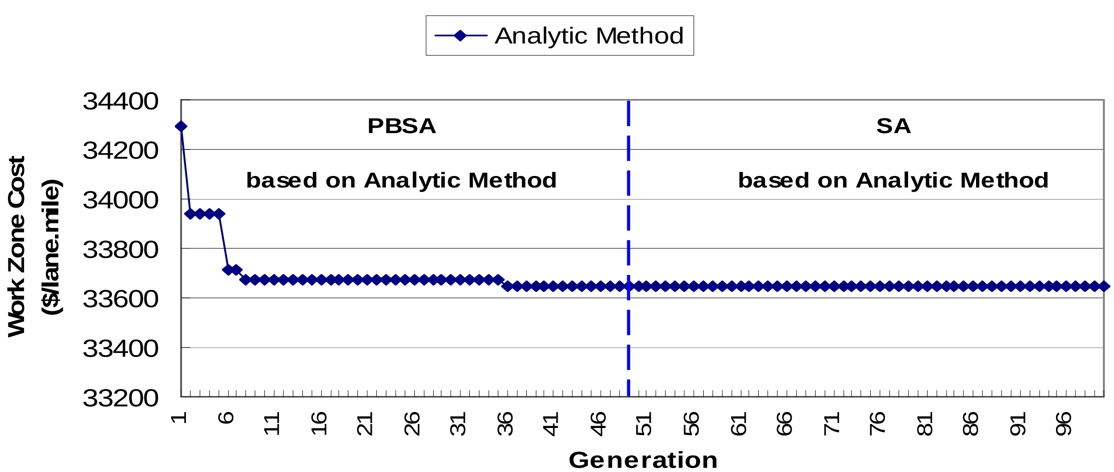


Figure 8.8 Optimization Convergence based on Analytic Method

8.3.3 Comparison of Optimized Results and Current Policies

The Maryland State Highway Administration (SHA) lane-closure policies for highway maintenance (Chen, 2003) are 9:00 a.m. -3:00 p.m. and 7:00 p.m. - 5:00 a.m. for single-lane closure; 10:00 p.m. -5:00 a. m. for two-lane closure; and 12:00 a.m. – 5:00 a.m. for three-lane closure.

In this numerical example, a single-lane closure is used. We compare the optimized results from the three optimization models with the current two policies. The work zone costs for all the five work zone plans are shown in Table 8.2 and Table 8.3. The comparisons are displayed in Figure 8.9.

Compared to the current policies based on either simulation model or analytic model, the optimized results yield lower total cost per lane mile. This comparison confirms that the optimization models developed in this study can significantly reduce the work zone total cost, including the agency costs and user costs.

Table 8.2 Comparison of Five Work Zone Plans based on Simulation Results

Simulation Results ($/lane.mile)	Optimized Results			Current Policies	
	Simulation method	Hybrid method	Analytic method	Daytime 6-hour	Nighttime 10-hour
Time Window	(16:00-4:00)	(16:00-4:00)	(16:00-5:00)	(9:00-15:00)	(17:00-5:00)
Agency Cost	33200	33200	33090	35000	33500
User Cost	101	101	226	208817	98
Total Cost	33301	33301	33316	243817	33598

Table 8.3 Comparison of Five Work Zone Plans based on Analytic Results

Analytic Results ($/lane.mile)	Optimized Results			Current Policies	
	Simulation method	Hybrid method	Analytic method	Daytime 6-hour	Nighttime 10-hour
Time Window	(16:00-4:00)	(16:00-4:00)	(16:00-5:00)	(9:00-15:00)	(17:00-5:00)
Agency Cost	33200	33200	33090	35000	33500
User Cost	521	521	558	211057	309
Total Cost	33721	33721	33648	246057	33809

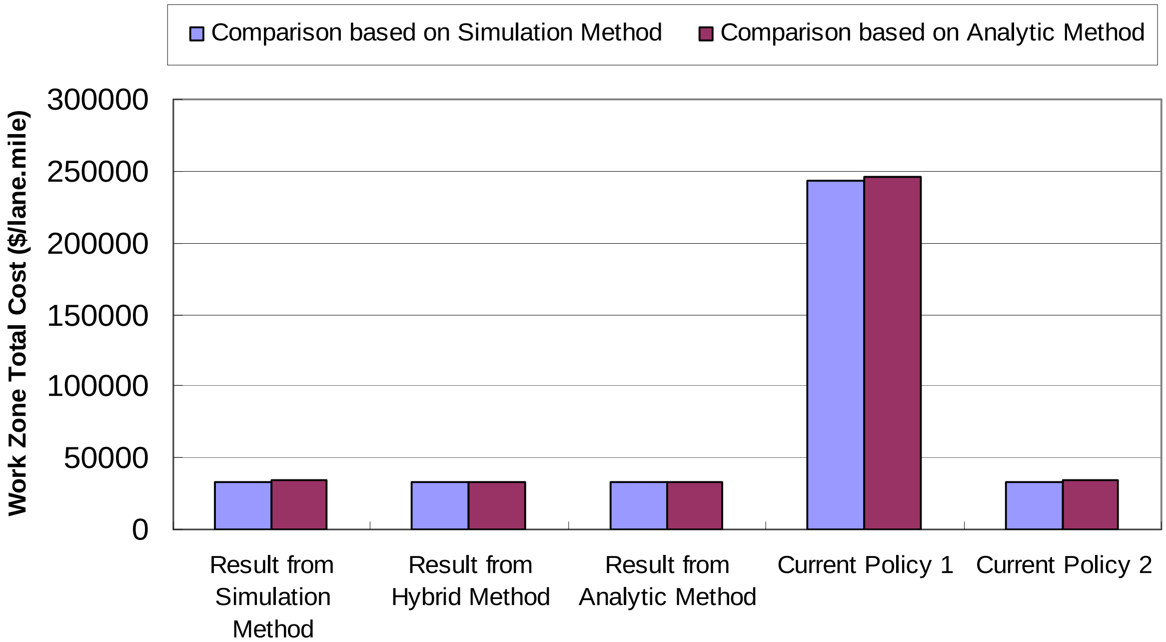


Figure 8.9 Comparison of the Optimized Results and Current Policies

8.4 Summary

In this chapter, an approach for optimizing work zone characteristics based on simulation is discussed. In this approach, the two-stage modified SA algorithm is applied to search an optimal solution while the CORSIM simulation is used to evaluate the objective function value.

To reduce the computational burden imposed by the simulation process, optimization through a hybrid method is proposed. The analytic method is used in solution evaluation in the first stage of the two-stage modified SA algorithm and the simulation method is applied in the second stage. While avoiding complete simulation in the early solution search phase, this hybrid method seek to combine the benefits of PBSA and SA algorithm as well as integrate the advantages of macroscopic analytic methods and microscopic simulation methods.

A numerical example is used to test three optimization models presented in the study: (1) optimization based on simulation method; (2) optimization based on analytic method; and (3) optimization through hybrid method. The three methods obtain similar optimized solutions, which reduce the work zone costs compared to current policies.

The optimization based on simulation is quite time-consuming compared to the other two methods. The SA algorithm works efficiently in the hybrid method, as we have expected.

Chapter 9 Conclusions and Recommendations

9.1 Project Summary

Our project aimed to develop a comprehensive work zone evaluation and decision support tool which can provide decision-makers with optimal solutions to lane closure strategies, scheduling, work zone configuration, traffic management and construction speed.

The project includes two phases. In Phase 1, CA4PRS (Construction Analysis for Pavement Rehabilitation Strategies), a scheduling and production analysis tool developed by researchers at University of California, Berkeley, is reviewed. CA4RPRS can provide users with useful answers to several "what-if" questions, but not optimized solutions. The work zone activities, construction procedures and work details considered in CA4PRS give us some references to modify the assumptions and formulations. Phase 2 was the major part of this project. In this phase, the following tasks have been completed:

(1) Based on our previous work, the analytic models for steady and time-dependent traffic inflows were modified to consider flexible lane closure strategies.

(2) To achieve more precise evaluation of work zone performance, the simulation model was introduced to simulate various work zone conditions and evaluate the user delay caused by work zone activities.

(3) Work zone optimization models based on the analytic method for steady traffic inflows, the analytic method for time-dependent inflows and the simulation method were presented. With the objective of combining the benefits of optimization and microscopic simulation, a hybrid approach was also developed to integrate the analytic method and the simulation method in the optimal search process.

(4) To integrate all the above works, a user-friendly software package has been developed.

9.2 Conclusions

In this project, four major types characteristics are considered to define a work zone activity: (1) lane closure alternatives including the number of closed lanes, the number of access lanes and the number of maintained lanes; (2) operation characteristics including the work zone length and work zone schedule; (3) work rate parameters including the zone setup time (z_1), the average maintenance time (z_2), the zone setup cost (z_3) and the average maintenance cost (z_4); (4) Detour types including no detour, single detour and multiple detours.

For two-lane two-way highways and multiple-lane two-way highways, work zone cost models, which can evaluate the agency cost, the user delay cost and the accident cost, are formulated. Three methods are developed to estimate the user delays caused by work zone activities: (1) the analytic method for steady traffic inflows; (2) the analytic method for time-dependent traffic inflows; and (3) the simulation method based on CORSIM, a comprehensive simulation program developed by Federal Highway Administration (1992).

The simulation model can provide detailed representations of traffic characteristics, network geometrics and traffic control plans. It is found that CORSIM estimates higher delays than the analytic method under uncongested traffic conditions, which is consistent with our expectation because a simulation process can record the acceleration delay, deceleration delay, shockwave delay and other components which are ignored in analytic methods. However, CORSIM estimates lower delays than the analytic method under very congested traffic conditions. This can be explained by the inability of CORSIM to calculate the delays of the vehicles that cannot enter the network as scheduled due to the queue spillbacks beyond traffic entry nodes in an over-saturated network.

Using the classic optimization techniques of differential calculus, an analytic optimization model is formed to jointly optimize the work zone length and work rates, considering construction time-cost tradeoff relation under steady traffic inflows. Sensitivity analysis shows that employing a more expensive but faster construction method can help minimize the total cost as the flow through the work zone increases. Unlike in previous studies, when the time-cost tradeoff is incorporated in the optimization, the optimal zone length L* does not always decrease as the inflow increases; that length is jointly optimized with work rate parameters based on several input parameters.

For time-varying traffic flows, another optimization method is developed to simultaneously optimize work scheduling, work zone length, lane closure alternative, traffic diversion fraction and construction speed, while considering the constraints on the lane closure times, lane closure length and maximum allowable diverted fractions. A heuristic algorithm, called the two-stage modified simulated annealing algorithm (2SA), is developed to find the optimal solution. In the first stage, an initial optimization process is performed through a population-base simulated annealing (PBSA) approach and the result will be sent to the second stage as a relatively good initial solution. In the second stage, a conventional simulated annealing (SA) method is used for refined search inside the promising region provided by the first step. A numerical experiment shows that PBSA performs better than traditional SA in solving the optimization problem with multiple local optima.

The objective function is to minimize the work zone cost. Based on the method used to evaluate the objective function value, the work zone optimization models can be classified into three categories:

(1) Optimization based on analytic method (A2SA), in which the analytic model for time-dependent traffic inflows is used to evaluate the objective function throughout the optimization process;

(2) Optimization based on simulation method (S2SA), in which we apply simulation to evaluate the objective throughout the optimization process;

(3) Optimization through a hybrid method (H2SA), in which the analytic method is used in the first stage to identify promising regions for solutions without fully simulating each solution, while the simulation method is applied in the second stage.

The methodologies have been tested in a hypothetical network, which can be analytically modeled. A detailed simulation model is also built for this network. Performing the optimization based on analytic method for different traffic levels, we find that as traffic inflows increase, fewer lanes can be allowed to close, the work zone length as well as the work zone duration decreases for any given lane closure alternative, a faster work rate is preferred and a higher diverted fraction is desirable. For the test case, three optimization models obtain similar solutions, which are better than current policies used by Maryland SHA for this test example. The simulation-based optimization takes much more time than the other two optimization models. The convergences of the optimization processes indicate that the SA algorithm in the second stage searches efficiently through the hybrid method. The hybrid method yields the same solution as the simulation-based optimization while greatly reducing the computation time.

9.2 Recommendations

In future studies, possible extensions of the analysis and models developed in this study are desirable, as follows:

1. Work Zone Cost Model

(1) When estimating the user cost, we consider the user delay cost and accident cost in this study. Currently, the accident cost is proportional to the user delays (veh.hr). It may be more realistic to also consider vehicle miles and the difference in safety costs for various time periods (for example, daytime may be safer than nighttime in work zones). Another significant component of road user cost, the vehicle operating cost, should be explicitly and separately estimated in future work.

(2) For the present work zone optimization, we assume the fixed setup cost (z_1) and its duration (z_3) are the same for all work zone configurations. It is also assumed the average maintenance time (z_2) and cost (z_4) can be affected by working on multiple lanes simultaneously. However, these time and cost parameters may vary with work zone configurations (for example, higher work zone setup cost and duration may be needed for crossovers), existence of access lanes (for example, more access lanes may increase work efficiency), time period of the work zone activity (for example, nighttime work zone may cost more than daytime work zone) and other factors. Therefore, the cost and duration parameters should be determined for different work zone characteristics.

2. Simulation model

In the current study, simulation is used to estimate the user delays caused by work zone activities. In fact, many other Measures of Effectiveness (MOE) can be obtained from the simulation outputs, such as density, speed, environmental effects and fuel consumption. We would seek to exploit more information provided by simulations in future research.

3. Extensions of Decision Variables

(1) Although we consider time-cost tradeoff in the optimization based on analytic method for steady traffic inflows, the work rates with different construction time-cost combinations are optimized discretely. By introducing a functional relation between variable cost and variable time per lane length, we can optimize the work rate in a continuous way. For reasonable simplicity and generality, the time-cost tradeoff function developed in Chapter 5

is assumed to be a hyperbolic one, which is convex and continuously differentiable. Actually, the time-cost relation is not restricted to a particular form. Further studies are desirable to consider more realistic cases, including piecewise linear functions.

(2) Besides work rate, work quality is also an important issue highway agencies care about. Work quality depends on work intention (e.g. the depth of the resurfacing work), materials (e.g. concrete or asphalt) and construction method (e.g. overlay or replacement). It not only affects the average work time and cost (z2 and z4) but also determines the new service life and future maintenance frequency. Decision variables related to work quality can be added in our optimization model. And in this case, costs per lane length per time unit would be used as measures of effectiveness.

(3) In this study, it is assumed a maintenance project is divided into recursive work zones with the same characteristics. For example, a project can be completed in four one-lane closure 8-hour nighttime work zones. In a future study, different work zones could be considered within one project. The transition and integration of zones should be considered if their configurations are different.

(4) In the current models the diversion fractions stay constant while one zone is resurfaced. However, diversion fractions which vary with time-dependent inflows may be considered for dynamic traffic control. Further research on time-depedent diversion fractions is desirable.

4. Consideration of Demand Changes

For long-term projects which are well publicized to motorists, users' travel behaviors may change. Therefore, traffic assignment considering user equilibrium may be needed to estimate the traffic as modified by work zone conditions.

5. Optimization Algorithm

In the hybrid approach proposed in this study, the analytic method and simulation method are used in different stages. Local search based on simulation method is performed in a relatively good neighborhood obtained in the first stage. The algorithm might be more efficient when employing both methods in the first stage using PBSA algorithm by performing multiple analytic optimization steps between each simulation step.

6. Parallel Computing

Although the use of heuristics can significantly reduce the temporal complexity of the search process, the latter remains time-consuming especially when objective functions are evaluated through simulation. Adapting the PBSA algorithm for parallel computing is a promising approach for speeding up the optimization process.

References

[1]. Al-Deek, H. M., Garib, A., and Radwan, A. E., (1995) “New Method for Estimating Freeway Incident Congestion,” Transportation Research Record 1494, Transportation Research Board, pp 30-39.

[2]. Bohachevsky, I. O. Generalized Simulation Annealing for Function Optimization. TECHNO-METRICS, 28(3):209-217, 1986.

[3]. Booz-Allen & Hamilton Inc. California Life-Cycle Benefit/Cost Analysis Model (Cal-B/C)—Technical Supplement to User's Guide. California Department of Transportation. September 1999b. Available at: http://www.dot.ca.gov/hq/tpp/planning_tools/tech_supp.pdf. Accessed Nov. 16, 2005

[4]. Booz-Allen & Hamilton Inc. California Life-Cycle Benefit/Cost Analysis Model (Cal-B/C)—User's Guide. California Department of Transportation. September 1999a. Available at: http://www.dot.ca.gov/hq/tpp/planning_tools/calbc_userguide.pdf. Accessed Nov. 16, 2005

[5]. Butcher, W. S. Dynamic Programming for Project Cost-Time Curves. Journal of the Construction Division, ASCE, Vol. 93, 1967, pp. 59-73.

[6]. CA4PRS Website. California Department of Transportation. http://www.dot.ca.gov/research/roadway/ca4prs/ca4prs.htm. Accessed July 20, 2005.

[7]. Cassidy, M. J. and L. D. Han. Predicting Vehicle Delays and Queue Lengths on Two-Lane Highways during Maintenance Activity. 71st Annual Meeting, Transportation Research Board, Washington, D.C. 1992.

[8]. Chen, C. H. and P. Schonfeld. Work Zone Lengths for a Four-Lane Road with an Alternate Route. Forthcoming in the Journal of Transportation Engineering, ASCE, 2005.

[9]. Chen, C. H. Intergrated Management of Highway Maintenance and Traffic, Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Maryland, College Park, December 2003.

[10]. Chen, C. H., P. Schonfeld, and J. Paracha. Work Zone Optimization for Two-Lane Highway Resurfacing Projects with an Alternate Route, 84th Annual Transportation Research Board Meeting (05-0561 on CD-ROM), forthcoming in Transportation Research Record, 2005.

[11]. Cheu, R. L., Y. Wang and T. F. Fwa . Genetic Algorithm-Simulation Methodology for Pavement Maintenance Scheduling. Computer-Aided Civil and Infrastructure Engineering, Vol. 19, 2004, pp. 446-455.

[12]. Chien, S. and P. Schonfeld. Optimal Work Zone Lengths for Four-lane Highways. Journal of Transportation Engineering, ASCE, Vol. 127, No. 2, 2001, pp. 124-131.

[13]. Chien, S., and P. Schonfeld, Optimal Work Zone Lengths For Four-Lane Highways. Journal of Transportation Engineering, ASCE, Vol. 127, No. 2, 2001, pp. 124-131.

[14]. Chien, S., D. G. Goulias, S. Yahalom and S. M. Chowdhury. Simulation-Based Estimates of Delay at Freeway Work Zones. Journal of Advanced Transportation, Vol. 36, No. 2, 2002, pp. 131-156.

[15]. Chien, S., Y. Tang, and P. Schonfeld. Optimizing Work Zone Lengths for Two-Lane Highway Maintenance Projects. Journal of Transportation Engineering, ASCE, Vol. 125, No. 1, 1999, pp. 21-29.

[16]. Chien, S., Y. Tang, and P. Schonfeld. Optimizing Work Zones For Two-Lane Highway Maintenance Projects. Journal of Transportation Engineering, ASCE, Vol. 128, No. 2, 2002, pp. 145-155.

[17]. Chitturi, M. V. and R. F. Benekohal. Comparison of QUEWZ, FRESIM and QuickZone with Field Data for Work Zones. In Transportation Research Board the 82th Annual Meeting. CD-ROM. Transportation Research Board, Washington, D.C., 2003.

[18]. CORSIM Run-Time Extension (RTE) Developer’s Guide. Federal Highway Administration. U.S. Department of Transportation, 2003.

[19]. Crowther, B. C. A Comparison of CORSIM and INTEGRATION for the Modeling of Stationary Bottlenecks. Master Thesis, the Virginia Polytechnic Institute and State University, 2001.

[20]. De, P., E. J. Dunne, J. B. Ghosh, and C. E. Wells. The Discrete Time-cost Tradeoff Problem Revisited. European Journal of Operational Research, Vol. 81, 1995, pp. 225-238.

[21]. Dixon, K. K., A. Lorschneider and J.E. Hummer. Computer Simulation of I-95 Lane Closures Using FRESIM. Papers for the

[22]. Dudek, C. L. and S. H. Richard. Traffic Capacity through Urban Freeway Work Zone in Texas. Transportation Research Record 869, Transportation Research Board, Washington, D.C., 1982, pp. 14-18.

[23]. E. B. Lee, J. T. Harvey, M. M. Samadian. Knowledge-based Scheduling Analysis Software for Highway Rehabilitation and Reconstruction Projects. The 84th Annual Transportation Research Board Meeting, Washington, D.C, 2005.
[24]. Feng, C.-W., L. Liu, and S. A. Burns. Using Genetic Algorithms to Solve Construction Time-Cost Trade-Off Problems. Journal of Computing in Civil Engineering, ASCE, Vol. 11, No. 3, 1997, pp. 184-189.
[25]. Fu, M. C. Feature Article: Optimization for simulation: Theory vs. Practice. INFORMS Journal on Computing. Vol. 14, No. 3, 2002, pp. 192-215.
[26]. Gerlough, D. L. and M. J. Huber. Traffic Flow Theory Special Report 165, Transportation Research Board, National Research Council, Washington, D.C., 1975, pp. 49-70.
[27]. Halati, A, H. Lieu and S. Walker. CORSIM-Corridor Traffic Simulation Model. Proceeding ASCE Conference on Traffic Congestion and Safety in the 21st Century, Chicago, 1997.
[28]. Jiang, X. and H. Adeli. Freeway Work Zone Traffic Delay and Cost Optimization Model. Journal of Transportation Engineering, ASCE, Vol. 129 No. 3, 2003, pp. 230-241.
[29]. Jiang, X., and H. Adeli. (2003). Freeway Work Zone Traffic Delay and Cost Optimization Model. Journal of Transportation Engineering, ASCE, Vol. 129, No. 3, 2003, pp. 230-241.
[30]. Jiang, Y., "A model for estimating excess user costs at highway work zones," Transportation Research Record, No.1657, 31-41 (1999)
[31]. Kang, K. P., G. L. Chang and N. Zhou. An Optimal Dynamic Speed Limit Control for Highway Work-Zone Operations. In Transportation Research Board the 83th Annual Meeting. CD-ROM. Transportation Research Board, Washington, D.C., 2004.
[32]. Kirkpatrick, s., C. d. Gelatt and M. P. Vecchi. Optimization by Simulated Annealing. Science, Vol. 220, 1983, pp. 671-680
[33]. Krammes, R. A. and G. O. Lopez. Updated Capacity Values for Short-term Freeway Work Zone Lane Closures. The 73rd Annual Transportation Research Board Meeting, Transportation Research Board, Washington, D.C, 1994.
[34]. Lee, E.B., D. Thomas and B. Loren. Planning Urban Highway Reconstruction with Traffic Demand Affected by Construction Schedule. Journal of Transportation Engineering, ASCE, Vol. 131, No. 10, 2005, pp. 752-761.
[35]. Lee. E.B., C. Kim and J. T. Harvey. Application of Macro- and Microscopic Simulations for the Traffic Planning of Urban Highway Reconstruction. In Transportation Research Board the 84th Annual Meeting. CD-ROM. Transportation Research Board, Washington, D.C., 2005.
[36]. McoCoy, P. T. and D. J. Mennenga. Optimum length of single-lane closures in work zones on rural four-lane freeways. Transp. Res. Rec. 1650, Transportation Research Board, Washington, D.C., 1998, pp. 55-61
[37]. Memmott, J. L. and C. L. Dudek. Queue and Cost Evaluation of Work Zones (QUEWZ). Transportation Research Record 979, Transportation Research Board, Washington, D.C., 1984, pp. 12-19.
[38]. Moder, J. J., C.R. Phillips, and E.W. Davis. Project Management with CPM, PERT and Precedence Diagramming. Van Nostrand Reinhold, New York, 1983.
[39]. Nash, S. G. and A. Sofer. Linear and Nonlinear Programming. The McGraw-Hill Companies, Inc., New York, 1996.
[40]. Pal, P. and K.C. Sinha. Evaluation of Crossover and Partial Lane Closure Strategies for Interstate Work Zones In Indiana. In Transportation Research Record 1529, TRB, National Research Council, Washington, D.C.,1996, pp.10-18.
[41]. Robinson, D. R. A Dynamic Programming Solution to Cost-Time Tradeoff for CPM. Management Science, Vol. 22, No. 2, 1975, pp. 158–166.
[42]. Rouphail, N. M. and G. Tiwari. Flow Characteristics at Freeway Lane Closures. Transportation Research Record 1035, Transportation Research Board, Washington, D.C., 1985, pp. 50-58.
[43]. Rouphail, N. M., Z. S. Yang and J. Fazio. Comparative Study of Short- and Long-Term Urban Freeway Work Zones. Transp. Res. Rec. 1163, Transportation Research Board, Washington, D.C., 1988, pp. 4-14.
[44]. Schonfeld, P. and S. Chien. Optimal Work Zone Lengths for Two-lane Highways. Journal of Transportation Engineering, ASCE, Vol. 125, No. 1, 1999, pp. 21-29.
[45]. Siemens, N. A Simple CPM Time-cost Tradeoff Algorithm. Management Science, Vol. 17, No. 6, 1971, pp. 354-363.
[46]. Steven, D. S. and T. H. Maze. Evaluation of Rural Interstate Work Zone Traffic Management Plans in Iowa Using Simulation. Mid-Continent Transportation Symposium, May 2000.

[47]. Sun, R.L, J. E. Dayhoff, J. E. and W. A. Weigand. A Population-Based Search from Genetic Algorithms through Thermodynamic Operation. Technical Research Report for the National Science Foundation (T.R. 94-78), 1994.
[48]. TSIS 5.1 User Manual. Federal Highway Administration. U.S. Department of Transportation, 2003.
[49]. Viera-Colon, S. I. Optimal Work Zones for Four-lane Highways. Master of Science Thesis, Department of Civil Engineering, University of Maryland, College Park, 1999.