2004-37

Final Report

Measuring the Equity And Efficiency of Ramp Meters

MINNESOTA DEPARTMENT OF TRANSPORTATION

Research

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.
MN/RC – 2004-37		

4. Title and Subtitle	5. Report Date
MEASURING THE EQUITY AND EFFICIENCY OF RAMP METERS	November 2004
	6.
7. Author(s)	8. Performing Organization Report No.
David Levinson, Lei Zhang, Shantanu Das, Atif Sheikh	
9. Performing Organization Name and Address	10. Project/Task/Work Unit No.
University of Minnesota Department of Civil Engineering 500 Pillsbury Drive S.E. Minneapolis, MN 55455-0116	11. Contract (C) or Grant (G) No. (c) 74708 (wo) 162
12. Sponsoring Organization Name and Address	13. Type of Report and Period Covered
Minnesota Department of Transportation Research Services Section 395 John Ireland Boulevard Mail Stop 330 St. Paul, Minnesota 55155	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

http://www.lrrb.org/PDF/200437.pdf

16. Abstract (Limit: 200 words)

Traffic congestion has become an increasingly serious problem in many cities. Ramp metering, which maintains smooth freeway mainline flow by limiting vehicle entry at entrance ramps, has been proposed and implemented in a number of metropolitan areas in and outside the U.S. to mitigate freeway congestion. This study aims to develop both efficient and equitable freeway ramp control strategies.

Traffic conditions with and without ramp metering are evaluated on several representative freeways in the Twin Cities with a comprehensive set of performance measures. A unified theory for ramp metering is proposed based on a linear programming model of freeway traffic dynamics. The most efficient ramp control algorithm is found to be also the least equitable one. A novel control objective, minimizing weighted or perceived travel time, is therefore proposed to balance efficiency and equity objectives of ramp metering. This research also develops a new family of applicable ramp metering strategies, which consider both efficiency and equity, and are demonstrated in a microscopic traffic simulator.

Future studies should compare various traffic control methods under the analytical framework proposed in this report. Researchers should also pursue field experiments of the proposed multi-objective ramp control strategies.

17. Document Analysis/Descriptors		18.Availability Statement
Ramp Metering Optimal Control Problem Co_EOAX	Freeway Operations Evaluation BEEX	No restrictions. Document available from: National Technical Information Services, Springfield, Virginia 22161

19. Security Class (this report)	20. Security Class (this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	102	

MEASURING THE EQUITY AND EFFICIENCY OF RAMP METERS

Final Report

Prepared by:
David Levinson
Lei Zhang (Research Assistant)
Shantanu Das (Research Assistant)
Atif Sheikh (Research Assistant)
Department of Civil Engineering
University of Minnesota

November 14

Published by:
Minnesota Department of Transportation
Office of Research Services
Mail Stop 330
John Ireland Boulevard
St. Paul, Minnesota 55155-1899

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Traffic congestion has become an increasingly serious problem in many cities. Ramp metering, which maintains smooth freeway mainline flow by limiting vehicle entry at entrance ramps, has been proposed and implemented in a number of metropolitan areas in and outside the U.S. to mitigate freeway congestion. The general purpose of this study is to develop both efficient and equitable freeway ramp control strategies.

First, we define measure of efficiency and equity for ramp meters, and apply them to data collected in the Twin Cities ramp metering shut-off experiment. The results of data analysis shows a trade-off between efficiency and equity. We developed a new time-dependent optimization program for the global optimal ramp control problem, which aims to minimize total travel time and does not require origin-destination demand information. We also provide a heuristic solution to the real-time version of the new optimization program, which can be directly implemented in real-world freeway networks. The heuristic provides a qualitative foundation for the development of an analytical framework for ramp metering under which to conduct future ramp metering studies.

In addition, we developed a new series of ramp control strategies, Co_EOAX with some additional equity considerations. Since the metering objective of minimizing absolute travel time is purely efficiency-oriented, we present a new objective for ramp metering, minimizing weighted (perceived) travel time, which has equity considerations built in when ramp delays are weighted non-linearly. To achieve this new metering objective, we developed a simulate method which is able to balance efficiency and equity of ramp control strategies. Co_EOAX was coded and tested in a microscopic traffic simulator and the optimal equity consideration in Co_EOAX was identified by the new metering objective.

Finally, we show in the simulator that the Co_EOAX with the optimal equity consideration is both more efficient and more equitable than the no-control scenario and an existing ramp control strategy.

This research provides both a unified theory for multi-objective freeway on-ramp traffic management, and a new family of applicable ramp metering strategies. Future studies should compare various traffic control methods under the analytical framework proposed in this report. Researchers should also pursue field experiments of the Co_EOAX strategies.

1 INTRODUCTION

Ramp metering has been applied since the 1960s in Chicago, IL; Detroit, MI; and Los Angeles, CA. The first application in Chicago in 1963 involved a police officer who would stop traffic on an entrance ramp and release vehicles one at a time at a predetermined rate, so that the objective of safer and smoother merging onto the freeway traffic was easier to achieve without disrupting the mainline flows (May 1964, Minnesota Department of Transportation (Mn/DOT) 2002a). The purpose of the ramp meters was to reduce congestion by managing traffic demand, to improve the efficiency of merging, and to reduce accidents. The success of these early applications led to the extension of ramp metering systems to many other metropolitan areas in the U.S. including Phoenix, AZ; Fresno, CA; Sacramento, CA; San Francisco, CA; San Diego, CA; Denver, CO; Atlanta, GA; Twin Cities, MN; Las Vegas, NV; Long Island, NY; New York, NY; Cleveland, OH; Lehigh Valley area, PA; Philadelphia, PA; Houston, TX; Arlington, VA; Milwaukee, WI; and Seattle, WA; as well as in cities in other countries including Sydney, Australia; Toronto, Canada; and Birmingham and Southampton, UK. Ramp meters have been withdrawn after initial introduction in several US cities, including Austin, TX; Dallas, TX; San Antonio, TX; and Columbus, OH.

Since the earliest practices of ramp metering, a number of reports have explored and documented various benefits of ramp meters (Pinnell et al. 1967, Newman *et al.* 1969, Fonda 1969, Gordon and Wynne 1973, Piotrowicz and Robinson 1995). Benefits attributed to ramp metering in the literature include lowering travel times, increasing freeway throughputs, reducing accident rates, reducing fuel consumption and emissions, and offering opportunities to provide priority entry to high-occupancy vehicles (HOV bypass). However, ramp meters have some negative effects, such as the impacts of diversion onto surface networks, and queuing on ramps. It has also been argued that ramp meters tend to reserve freeways for commuters traveling from the suburbs while discouraging urban users, which is a serious equity debate. Despite the fact that the literature acknowledged these negative impacts as early as in the 1960s (Pinnell *et al.*

1967), many ramp metering evaluation studies focus only on the beneficial side of ramp metering and fail to fully analyze negative effects.

An equity measure is able to capture the aforementioned negative impacts of ramp meters. The major purpose of this study is to develop both an efficient and an equitable ramp metering control strategy. To achieve this, one must explicitly define what efficiency and equity mean for ramp meters.

The study will first develop a comprehensive set of performance measures to evaluate ramp metering systems including mobility and equity, as well as productivity, consumers' surplus, accessibility, travel-time variation and travel-demand responses. These measures of effectiveness (MOEs) will then be applied to evaluate an existing ramp metering control strategy, the Minnesota Zonal Algorithm (Minnesota algorithm), based on a before/after data set collected from the 2000 Twin Cities ramp metering shut-off experiment. It will be kept in mind in developing the evaluation methodology that the computation of various MOEs only requires data that is routinely collected by the Mn/DOT Traffic Management Center (TMC). This differs from other studies using freeway simulation models which usually take advantage of outputs from simulators that are not measurable by installed hardware in field (Reiss *et al.* 1991, Stephanedes *et al.* 1992). Therefore, using these MOEs, studies can evaluate freeway system performance routinely without any additional data collection efforts. The impacts of continuous traffic growth or changes to control strategies on the freeway system performance hence can be tracked overtime.

In this freeway performance evaluation, we give special attention to the consistency of the evaluation results from various MOEs. If, after considering multiple measures, we draw the same conclusion (e.g., ramp control under the Minnesota algorithm is better than no control), the evaluation results will be very reliable. But if different conclusions are drawn from different MOEs, there will be no clear answer, but rather trade-offs. Some previous studies have observed a trade-off between mobility and equity in implementing specific ramp control strategies (e.g. Kotsialos and Papageorgiou 2001). The effort to identify potential trade-offs among various performance measures in this report is more comprehensive and quantitative.

The major portion of this report will be devoted to the development of a new efficient and equitable ramp metering control strategy. It starts with the formulation of a new globally optimal time-dependent ramp control problem which minimizes the total freeway system (freeway mainline and ramps) travel time. By recognizing that one may not need to know detailed time-sliced origin-destination (OD) information to achieve optimal control, this new formulation only requires input data that are all directly measurable in real-time. This is a significant improvement over the previously-proposed global optimal ramp control formulations (Wattleworth 1963, Wattleworth 1967, Papageorgiou 1995, M. Zhang et al. 1996, M. Zhang and Recker, 1999, Lovell 1997, Lovell and Daganzo, 2000, Chang et al. 2002) in that the difficulties of obtaining accurate OD information in the previous formulations are successfully avoided. We also develop a real-time (rolling-horizon) version of the new formulation, and provide a heuristic solution to it. The heuristic solution is shown to be the globally optimal solution on a small-scale network and is very likely to be the globally optimal solution to larger networks when certain conditions hold.

The heuristic solution to the new formulation also makes qualitative sense. An analytical framework for ramp metering control will then be built based on it, with the hope of leading towards a more unified and generic ramp control theory. Under this framework, it is possible to decompose and study various individual elements that constitute a ramp control strategy individually, and hence many existing and/or proposed ramp metering strategies become more comparable.

A new ramp metering control strategy – **E**fficiency **O**riented **A**lgorithm (**EOA**), which is the direct implementation of the heuristic solution, will be coded and tested in a microscopic traffic simulator – AIMSUN2. The theory suggests that EOA is the most efficient ramp control strategy (where efficiency is measured by total travel time) but also the least equitable one. Minimizing total absolute travel time thus is not a “good” ramp control objective because the derived control strategy is not acceptable to society. To obtain a both efficient and equitable ramp control algorithm, we introduce the idea of minimizing total weighted travel time in which delayed time has higher weights than free flow travel time and longer delays have even higher weights than shorter delays.

Formulating a freeway system globally optimal program minimizing total weighted travel time turns out to be a quite difficult problem. Therefore, instead of minimizing weighted travel time directly, a series of ramp metering control strategies – **Co**ordinated **E**fficiency **O**riented **A**lgorithms with global grouping factor **X** (**Co_EOAX,** X = 1, 2, …) will be derived from EOA with different degrees of equity considerations. All these Co_EOAXs will then be coded and simulated in the AIMSUN2 microscopic traffic simulator and the resulting total weighted travel time will be computed and compared. The strategy that gives the minimum total weight travel time – Co_EOAX_{opt} ($_{opt}$: optimal), should be able to give us an appropriate balance between efficiency and equity. Finally, the relative efficiency and equity of the Minnesota algorithm and the Co_EOAX_{opt} will be compared using the MOEs developed in the early part of the report.

The report will be organized into eight chapters. Chapter 2 covers details of the development of various MOEs and the evaluation results of the Minnesota algorithm using these MOEs. Chapter 3 presents the new formulation of the globally optimal ramp control problem minimizing total travel time and the heuristic solution process. In Chapter 4, a framework for ramp metering studies is introduced, followed by detailed description of EOA and Co_EOAXs, implementations of the heuristic solution with different degree of equity considerations in Chapter 5. Chapter 6 explains the methodology of using microscopic simulator to identify the new strategy that minimizes the total weighted travel time. Simulation results, as well as the comparison between the new algorithm and the Minnesota algorithm, are presented in Chapter 7. Finally, summary remarks and conclusions are offered in Chapter 8.

The remainder of this introductory section will include a review of both the theoretical development and practices of ramp metering and some background information that is related to the discussion herein, including the Twin Cities ramp metering shut-off experiment, the Minnesota algorithm and the AIMSUN2 microscopic traffic simulator.

1.1 Literature Review of Ramp Metering Theories and Practices

The main goal of this section is to summarize the previous theoretical studies and practices that are most closely related to the discussions in this study. More comprehensive reviews are cited for interested readers.

1.1.1 Local Control

The earliest ramp metering control in the 1960s is pre-timed control in which ramp metering rates are fixed values based on historical data and engineering judgment. Local traffic responsive control was also first implemented in the 1960s. A local traffic responsive ramp metering control is developed based on real-time traffic information collected in the vicinity of individual on-ramps. The local traffic information of a ramp is then provided to a controller and the controller determines the actual ramp metering rate of this particular ramp via some feedback rules. Theoretical studies on local control focus on the effectiveness of various types of controllers, e.g. linear controller (Papageorgiou et al. 1991), Artificial Neuron Network controller (M. Zhang 1997) and fuzzy-logic controller (Taylor *et al.* 1998).

1.1.2 System-wide Optimization

The very first attempt to solve the ramp metering control problem via optimization at the freeway system level goes back to Wattleworth in two papers (Wattleworth, 1963, Wattleworth, 1967). The linear program proposed in these papers is essentially a time-invariant optimization program which aims to maximize the total input to the system at all on-ramps constrained by a set of restrictions. This objective can be shown equal to minimizing the total travel time in the freeway system under certain conditions. Several other optimization objectives were also proposed (see Lovell, 1997 for a review). Some distinct natures of these models include:

- They all incorporate a constraint equation which ensures that freeways are operated under free-flowing conditions. Hence, the difficulties of dealing with freeway mainline dynamics are avoided.
- They assume OD demand information is available.

- They assume that there are no diversions from the freeway system to surface arterial streets.

Recently, Lovell and Daganzo (Lovell and Daganzo 2000) extended Wattleworth's steady-state model to include time-dependency. Their formulation of a continuous version of the globally optimal time-dependent ramp control problem is not directly solvable, but a computationally efficient greedy heuristic solution is developed. However, the heuristic solution is only appropriate for small-scale network and OD information is still a required input.

There is also a body of literature combining the idea of the optimal control theory and macroscopic traffic flow models (Papageorgiou 1995, Kotsialos et al. 2002, M. Zhang et al., 1996, M. Zhang and Recker, 1999, Chang et al. 2002). Models of this type usually try to minimize total freeway system travel time directly (corridor travel time in Kotsialos et al. 2002). The freeway mainline dynamics are then described by a set of time-discrete equations based on finite difference approximation of specific macroscopic traffic flow models. However, this array of ramp metering control strategies also confront the difficulty of getting accurate OD information. Recently, some studies attempt to combine some OD estimation techniques with these strategies (e.g. Chang et al. 2002). Two drawbacks prevent these strategies from being implemented: the unclear reliability of the predicted OD information, and the computational complexity. Furthermore, it is still an unsolved problem to determine the globally optimal solutions for this type of system optimization programs. The effectiveness of such strategies can be determined only by simulating or implementing some locally optimal solutions. Therefore, although these studies developed methodologies to control ramp meters, their practical implications are limited.

1.1.3 Ramp Metering Practices

Numerous practical ramp metering algorithms have been developed since the 1960s. The first paragraph of this chapter gives a list of cities with ramp meters deployed. Each of these cities has its own ramp metering algorithm. Beside algorithms currently being operated, there are also many proposed algorithms, e.g. BALL Aerospace/FHWA (Federal Highway Administration) algorithm, ARMS (Advanced Real-time Metering System), ANN (Artificial Neural Network), MIT (a dynamic metering control module developed in Massachusetts Institute of Technology), and online simulation. A summary of a major portion of these algorithms is given by Bogenberger and May (1999). Some comparison studies among these algorithms have also been completed, mostly in simulators (Kwon et al. 2000, Lomax and Schrank 2000). However, the usefulness of the results from these comparison studies are quite limited largely because the studies failed to answer why one particular algorithm is better than the other(s). This happens because there are many elements in each algorithm, such as degrees of ramp coordination, threshold values and control variables, and each of these elements can affect the performance of the whole algorithm. Therefore, an analytical framework under which these individual factors can be easily decomposed is in order.

1.2 Background

This section briefly describes some background information that is related to this study, including the Twin Cities ramp metering shut-off experiment, the Minnesota algorithm and the AIMSUN2 microscopic simulator.

1.2.1 Twin Cities Ramp Metering Shut-off Experiment

The Minnesota Department of Transportation first installed meters in 1969 on entrance ramps to I-35E at Maryland Avenue and Wheelock Parkway in St. Paul. Additional ramps were brought into service on the same freeway (Mn/DOT 2002a). Observed improvements in traffic flow, higher speeds and fewer accidents led to more ramp meters in 1971. In Minneapolis, metering began on I-35W through south Minneapolis in 1970, followed by meters near the Lowry Hill tunnel on I-94. In 1972, Mn/DOT built a Traffic Management Center (TMC) in downtown Minneapolis to house the equipment and staff necessary for expanded metering. Over time, metering was added on nearly every entrance to the I-494/I-694 beltway, except on the metro area's east side, and on most highways inside the I-494/I-694 ring. Figure 1-1 shows Twin Cities ramp meter installations from inception. We can clearly see that there have been sharp increases in their numbers in some years. In the 1990s Mn/DOT adopted a policy favoring metering over capacity expansion in the Twin Cities area, primarily on a cost-effectiveness basis, as well as the belief that meters were easier to implement politically than new roads. By 2000, there were over 443 meters, mostly controlled in real-time by the TMC. Also, initially there were single lane ramp meters, but to avoid spillback onto arterial or connecting roads, the usual practice now is to have two-lane ramp meters.

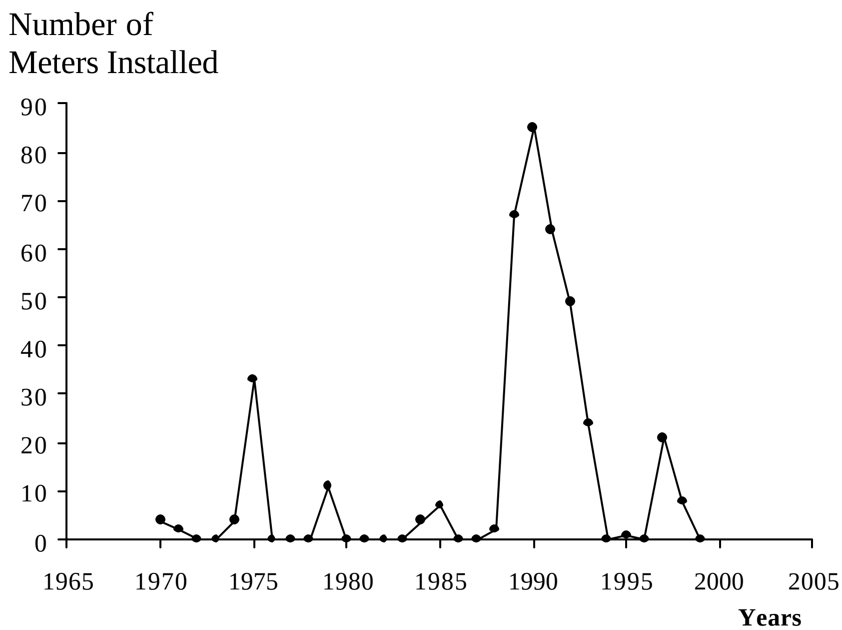


Figure 1-1 Ramp Meter Installation Timeline in the Twin Cities

Ramp meters began to assume an increasingly important role in the freeway management of the Twin Cities. Since 1970, Mn/DOT has spent nearly $4 million to install and upgrade ramp meters, television surveillance and other traffic management measures (Mn/DOT 2002a). Initially, meter installation did not draw political attention or spark public policy debates because it was seen to have helped smooth congestion without major inconvenience to drivers. Only in recent years, with congestion increasing and meter waits stretching beyond the level of driver acceptability (some waits in excess of 20 minutes), have drivers and legislators begun to question the meters' value, believing at best that the system was inefficiently managed by Mn/DOT.

The most vocal of metering opponents was Minnesota's Senate Republican leader Dick Day (Owatonna), who argued that the ramp meters cause both congestion and driver frustration. Early in 1999, Senator Day proposed turning off the ramp meters for a temporary period to test their effectiveness. In October 1999, he put forward his "Freedom to Drive" plan which promised to ease metro area congestion by turning off the ramp meters, raising speed limits, and using the left lane only for passing (Star Tribune 1999). The proposal was passed in the Senate but failed in the House in the 1999 legislative session. Senator Day argued that Mn/DOT had been withholding information from the public about the real effectiveness of ramp meters and had exaggerated the advantages of metering in order to continue funding metering projects while ignoring frustrated commuters waiting on the ramps. He linked his passion for turning off ramp meters to Americans' love for their cars, of being independent and driving unaffected by "unnecessary waits" on the ramps. Mn/DOT's support of delay-inducing metering was linked to support for new rail transit lines (the financial support of the Hiawatha Light Rail line connecting downtown Minneapolis with the Airport and Mall of America was an active question at the time – it has since passed).

The Transportation Finance Bill passed both houses of the state legislature in May 2000, increasing spending on roads and bridges, and including a provision for a month long "Ramp meter holiday" slated to take place in October 2000. Mn/DOT extended the study to two months. The bill required a study of the effectiveness of ramp meters in the Twin Cities metro area by conducting a shutdown study before the next legislative session.

1.2.2 Minnesota Zonal Algorithm[1]

The Minnesota Department of Transportation operated ramp meters before the shut-off experiment under a zonal control strategy that divided freeways into zones

[1] The Minnesota Zonal Algorithm is the ramp control strategy operated by the Minnesota Department of Transportation before the 2000 ramp metering shut-off experiment. A new Stratified Zonal Algorithm was developed after the experiment. Throughout the rest of this report, "the Minnesota algorithm" refers to the Minnesota Zonal Algorithm pre October, 2000.

terminating at bottlenecks. The number of vehicles in each zone at a given time was constrained by the capacity of the bottleneck. Ramp metering was used to limit those vehicles. The metering zone equation can be expressed as:

$$M + F = X + B + S - A - U \quad (1\text{-}1)$$

s.t. incident override and occupancy control

Where:

M sum of metered local access ramp flows (controlled variable);

F sum of metered freeway to freeway access ramp flows (controlled variable);

X sum of exit ramp flows (measured variables);

B downstream bottleneck flow at capacity (constant);

S space available within the zone (flow based on a measured variable).

A upstream mainline flow (measured variable);

U sum of unmetered entrance ramp flows (measured variable);

Any measured variation in $(X + B + S - A - U)$ is equaled by a controlled variation in $(M + F)$. Each individual variable, except *S* (set to be zero indicating no available space in the zone), in the zone equation (1-1) is assigned a 1-hour flow derived from historical detector data. When these historical flows are placed in equation (1-1) an exact balance may not appear. For this reason, a minor adjustment to the incoming flow *(A)* is made to balance the new equation and the values in this balanced equation (1-2) are the *target* flows:

$$M_T + F_T = X_T + B_T + S_T - A_T - U_T \quad (1\text{-}2)$$

Where:

$_T$ means *target* flows;

S_T set to zero.

Each metered ramp is assigned six metering rate factors. On local access ramps, these rates over a 5-min time period range from 0.5 to 1.5 (0.5, 0.7, 0.9, 1.1, 1.3 and 1.5) and on freeway-to-freeway ramps, from 0.75 to 1.25 (0.75, 0.85, 0.95, 1.05, 1.15, 1.25).

The narrower range for freeway-to-freeway ramps reflects the fact that these ramps carry much higher flows and a smaller percentage change provides the desired numerical flow change. The control variable $(M + F)$ is then expressed as the products of the target flows and metering rate factors. For example, at the most restrictive rate factor (rate 6), $M = 0.5M_T$ and $F = 0.75F_T$. The selection of which one of the six rate factors to use for a metering zone is then determined by a comparison of the on-line measured variables $(X + B + S – A – U)$ to a series of thresholds. For instance, if $0.6M_T + 0.8F_T < (X + B + S – A – U) < 0.8M_T + 0.9F_T$, rate 5 will be used. At rate 5, $M = 0.7M_T$ and $F = 0.85F_T$.

Whether any rate is the final metering rate factor is still subject to incident override and occupancy control. For more details on the Minnesota algorithm, readers may refer to Mn/DOT (1996, 1998) and Bogenberger and May (1999).

1.2.3 AIMSUN2 Simulator

AIMSUN2 (**A**dvanced **I**nteractive **M**icroscopic **S**imulator for **U**rban and **N**on-urban **N**etworks) is a microscopic traffic simulation package which can model various traffic networks. It can also be used for simulating adaptive control systems including ramp meters. The input required by AIMSUN2 is composed of three types of data: network description, traffic demand data and traffic control plans. The outputs provided by AIMSUN2 are a continuous graphical representation of the traffic network performance, statistical data (flows, speed, travel times, delays, total travel, etc.) and data gathered by the simulated detectors (counts, occupancy, etc.). AIMSUN2 is integrated in GETRAM (Generic Environment for Traffic Analysis and Modeling) (Barceló et al. 1994), a simulation environment comprising a traffic network graphical editor called TEDI, a network database, and an API (Applications Programming Interface). The API provides an interface for users to both read real-time simulation data from the AIMSUN2 simulator and pass control parameters generated by their own traffic control plans back to the simulator. Researchers in the University of Minnesota have further developed a CPI (Control Plan Interface) which contains some functions that are particularly helpful for coding third party ramp metering control strategies that work with the AIMSUN2 simulator (Hourdakis and Michalopoulos 2002, Muralidhar 2001). In Chapter 6 and 7 of this report, the AIMSUN2 simulator and the CPI will be used to simulate the coded new ramp control strategies.

2 EVALUATING RAMP METERS

Ramp meters in the Twin Cities have been the subject of a recent test of their effectiveness, involving turning them off for 8 weeks (see section 1.2.1). This chapter analyzes the results with and without ramp metering for several representative freeways during the afternoon peak period. Seven performance measures: mobility, equity, productivity, consumers' surplus, accessibility, travel time variation and travel demand responses are compared. It is found that ramp meters are particularly helpful for long trips relative to short trips. The trade-off between mobility and spatial equity is evident. The results are mixed, suggesting a more refined ramp control algorithm, which explicitly considers ramp delay, is in order.

The first section of this chapter outlines the various performance measures and shows how they are computed. The following section details the method used to measure travel times at on-ramps and freeway mainline segments from the data available. Then the data used in this study is introduced, followed by the results on studied freeways for each of the performance measures. The implications of these evaluation results on the development of new ramp metering control strategies are delivered at the end of this chapter.

2.1 Performance Measures

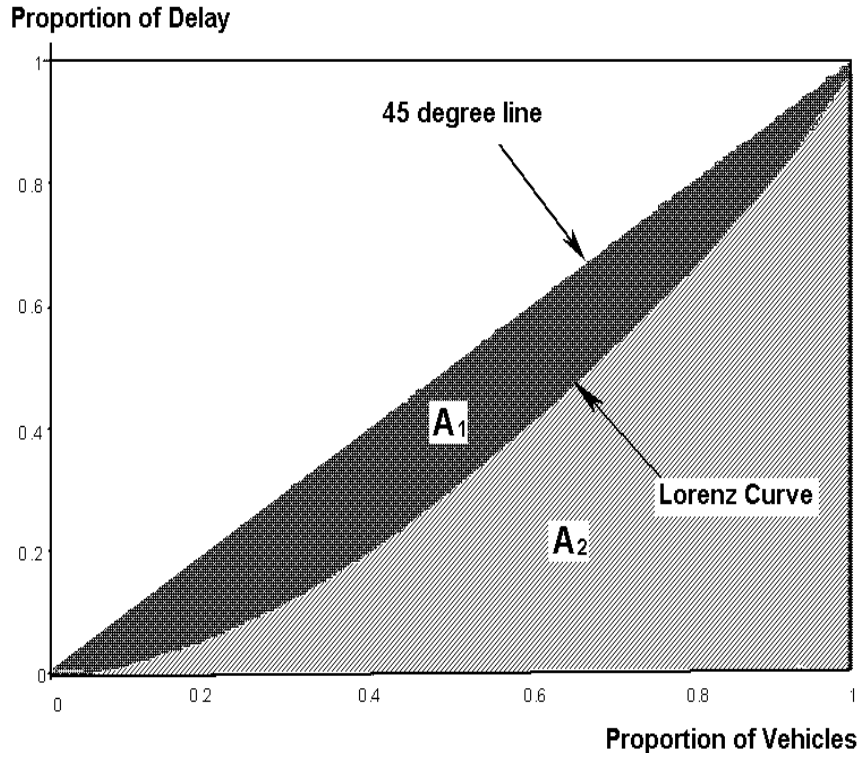
The evaluation of transportation systems garners significant attention in the planning, engineering, policy, management, and economics literatures, each of which approaches the problem differently, with unique concerns and objectives (Levinson 2002a,b). Since each field measures something different, it is important to consider each of those measures when evaluating the ramp metering system. If, after considering multiple measures, we draw the same conclusion (e.g. metering is better than no metering), we can be confident of our result. But if we draw different conclusions, there is no clear answer, but rather trade-offs. This section intentionally outlines a plurality of measures, reflecting the mobility of traffic engineers, the consumers' surplus of economists, the productivity of managers, the accessibility of planners, as well as the equity of policy analysts. We also consider demand changes and reliability, which are emerging issues for all of the fields.

The equity measure is particularly important to understand the differing perceptions of ramp meter effectiveness. Who, when, and where benefited and lost as a result of meters? Just because society gains does not mean everyone does. While system mobility was in the minds of the traffic engineers who designed and administered the system, individual travel times and their reliability were in the minds of each traveler and legislator who used the system. Accessibility has entered the public debate peripherally through the question of whether ramp meters cause sprawl – is access enhanced for suburban areas at the expense of the central city and first ring suburbs (Wascoe 2000, Zhang, 2002). We look at both overall accessibility as well as spatial equity to answer these questions. The question of consumers' surplus matters, though it has not entered the debate, because it is important to consider that some travelers are metered off the roads so that others can travel faster, these changes in demand change overall welfare. Similarly productivity lets us consider whether we are getting more output from the system per unit input, and is particularly important to commercial users of the system.

2.1.1 Mobility and Equity

Mobility, or ease of movement can be measured at on-ramps, freeway segments and for the system as a whole. For mobility, elements in two Origin-Destination (OD) trip mobility matrices (travel time, speed) can be averaged at various levels, which enable us to look at mobility of different groups of freeway users (e.g. drivers who have the same origin and destination; drivers who enter the freeway in the same time interval; all drivers during the peak period).

The Gini Coefficient of concentration (Gini 1936) and the Lorenz Curve (Lorenz 1905) have been used in economic studies to analyze income inequality. They are applied here to analyze inequalities of travel time and speed. Using ramp delay inequality as an example (Figure 2-1 gives the Lorenz Curve for the case of the Valley View Road on-ramp on TH169 northbound during an afternoon peak period), the Lorenz curve relates



A_1 area between Lorenz Curve and 45-degree line;
A_2 the rest of triangle defined by Lorenz Curve, X-axis, and vertical line projecting from 100% of the population in question.

Figure 2-1 Lorenz Curve for Valley View Road On-Ramp on TH169

the proportion of the population receiving a given proportion of delay. While the bottom 100% of the population gets 100% of delay by definition, the luckiest 50% may only get 30% of the delay. Then the Gini coefficient corresponding to this Lorenz curve can be computed, which is a numerical scalar indicating the degree of ramp delay inequality. Numerically, the Gini coefficient is equal to the ratio of $A_1/(A_1 + A_2)$ in the graph. A much easier way to obtain the Gini coefficient for the ramp delays experienced by a group of drivers $(d_1, d_2, \ldots, d_V)$, without constructing the Lorenz Curve, is to compute the "relative mean difference", i.e., the mean of the delay difference between every possible pair of drivers, divided by the mean delay (Glasser, 1962) :

$$G = \frac{\sum_{v=1}^{V}\sum_{u=1}^{V}|d_v - d_u|}{2V\sum_{v=1}^{V} d_v} \tag{2-1}$$

Where:

G Gini coefficient;

V total number of drivers;

d_v ramp delay experienced by driver v.

A Gini Coefficient of 0 indicates perfect equality, and 1 indicates perfect inequality (one driver suffers all the delay).

For each OD trip mobility matrix, spatial equity, temporal equity, and spatial-temporal equity are evaluated. The computation of these three equity measures and corresponding mobility measures is illustrated using an OD trip speed matrix (see Table 2-1). If a Gini coefficient is calculated using speeds in all time intervals for a specific OD pair (e.g. *speed 11*, *speed 21*, … , *speed K1* for OD pair 1), this Gini coefficient reflects temporal equity of this specific OD pair and is a temporal equity measure. Similarly, the average of these speeds is a temporal mobility measure. Temporal Gini coefficients measure the equity among drivers whose trips have the same origin and destination but start at the origin (an entrance ramp) in different time intervals. On the other hand, if a Gini coefficient is computed using speeds of trips that start in the same time interval but with different OD pairs (e.g. *speed 11*, *speed 12*, … , *speed 1P*), this Gini coefficient evaluates spatial equity in this time interval and thus is a spatial equity measure. The

Table 2-1 Speed OD Matrix

Speed	OD pair *1*	OD pair *2*	…	OD pair *P*
Time Interval *1*	*Speed 11*	*Speed 12*	…	*Speed 1P*
Time Interval *2*	*Speed 21*	*Speed 22*	…	*Speed 2P*
…	…	…	…	…
Time Interval *K*	*Speed K1*	*Speed K2*	…	*Speed KP*

average of these speeds is a spatial mobility measure. Finally, to evaluate the equity and mobility of the whole freeway network throughout a peak period, the spatial-temporal equity and mobility measures can be used respectively, which are calculated by all the $m \times n$ speed values.

2.1.2 Consumers' Surplus

Consumer surplus measures the difference between the price of a particular good and the reservation price the buyer is willing to pay for that good. Because the willingness to pay varies between individuals, the consumer surplus also varies. However, consumer surplus can be aggregated across individuals to get consumers' surplus. In general, change in consumers' surplus is used rather than absolute consumers' surplus, because it can be more readily measured. In travel, the price is the travel time while the quantity of the good is the number of vehicles using the road network. As the travel time across a certain stretch of the traffic network decreases, there will be more vehicles willing to access the same stretch pushing the flow up. Mathematically, this aggregate change in consumers' surplus (ΔCS) can be approximated as (the shaded area in Figure 2-2):

$$\Delta CS = 0.5\left(Q_{off} + Q_{on}\right)\left(t_{off} - t_{on}\right) \qquad (2\text{-}2)$$

Where:

Q_{on}, Q_{off} flows when the ramp meters are on, off respectively;

t_{on}, t_{off} travel times when the ramp meters are on, off respectively.

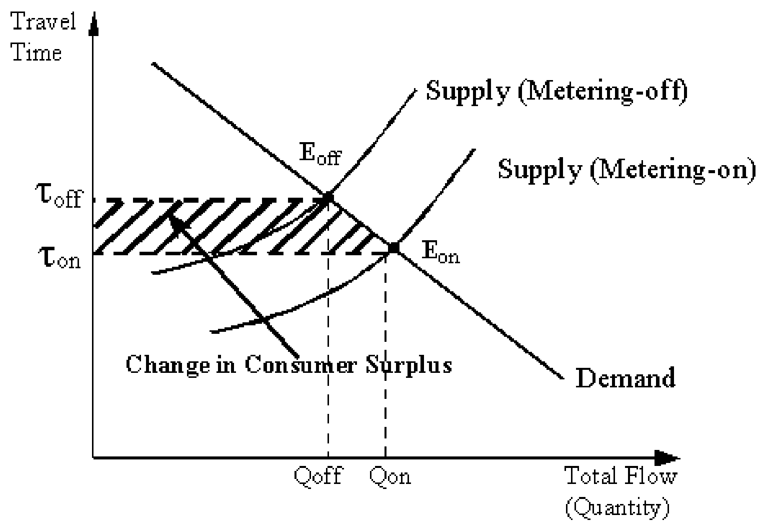
If the travel time when the ramp meters are switched off is greater than when they are switched on, then there is an increase of consumers' surplus with metering.

2.1.3 Productivity

Productivity is the ratio of the output of any product to the input that was required to produce that output. For transportation networks, vehicle kilometers traveled (*VKT*) and the vehicle hours traveled (*VHT*) are the output and the input respectively, equivalent to a measure of average speed.

$$P = \frac{output}{input} = \frac{\sum VKT}{\sum VHT} = \frac{\sum QL}{\sum Qt} \tag{2-3}$$

The ratio of *VKT* and *VHT* is measured for each freeway segment and ramps separately and then added to obtain the productivity of the system for both the metering-on and metering-off cases. Equation (2-3) gives a partial-factor productivity measure since *VKT* is not the only output of a freeway system and *VHT* is not the only input.



In equation (2), we use a straight line to approximate the actual demand curve between two network equilibrium points: E_{on} and E_{off}.

Figure 2-2 Measuring Changes in Consumers' Surplus

2.1.4 Accessibility

From the transportation planners' perspective, accessibility is defined as the ease of reaching destinations. Weibull (1980) suggests that accessibility is a measure of an individual's ability to participate in activities in the environment. Areas with high congestion often have high accessibility (Levinson and Kumar 1997). Three functional forms for the travel time decay function are used which are described in the results section. Freeway accessibility is computed as:

$$A_i = \sum_j Q_j f(t_{ij}) \tag{2-4}$$

where:

A_i accessibility of zone i;

Q_j opportunities at zone j, measured using exit flows;

t_{ij} travel time between i and j.

2.1.5 Travel Time Variation

Travel time variation is depicted as the standard deviation of travel times, and used as a measure of travel time reliability. Here we look at two ways to calculate travel time variation: inter-day and intra-day (Bates et al. 1987). Equation (2-5) is used to calculate inter-day travel time deviation of all trips with the same OD which start at the same time interval across different days. V_t stands for the *inter-day travel time variation* of this particular OD pair in time interval t.

$$V_t = std(t_{t,1}, t_{t,2}, \ \ ..., \ \ t_{t,n}) \quad (2\text{-}5)$$

where:

V_t *Inter-day travel time variation* of trips starting at time interval t;

std(.) The standard deviation of *(.)*;

$t_{t,n}$ Travel time of trips starting at time interval t in day n.

The inter-day travel time variation difference (D_t) between the metering-on case ($V_{t,on}$) and the metering-off case ($V_{t,off}$) can be obtained by simply subtracting metering-on values from metering-off values:

$$D_t = V_{t,off} - V_{t,on} \quad (2\text{-}6)$$

Then for trips with the same OD pair, we acquire a vector of inter-day travel time variation differences across different time intervals: D_1, D_2, … , D_t. By repeating this process, we acquire the same type of vectors for all OD pairs. The ranges (lower/upper bound) and median values of those vectors can be illustrated graphically with a range/median plot (shown in the result section).

On the other hand, v_n is the *intra-day travel time variation* of all afternoon peak trips (in different time intervals) with the same OD in day n only. The equation for calculating v_n is:

$$v_n = std(t_{1,n}, t_{2,n}, \ \ ..., \ \ t_{t,n}) \quad (2\text{-}7)$$

where:

v_n Intra-day travel time variation of trips in day *n*.

The averages of intra-day travel time variation over all days are first computed separately for metering-on and metering-off cases. Then by comparing these two average values ($\bar{v}_{on}$ and $\bar{v}_{off}$, equation (2-8)) for each OD pair, we ascertain whether ramp meters reduce or increase intra-day travel time variation.

$$\begin{aligned} \bar{v}_{on} &= average(v_{1,on}, v_{2,on}, ..., v_{n,on}) \\ \bar{v}_{off} &= average(v_{1,off}, v_{2,off}, ..., v_{n,off}) \end{aligned} \quad (2\text{-}8)$$

where:

average(.) The average of *(.)*;

2.1.6 Demand Responses

Demand responses may not be counted as a measure of effectiveness in a rigorous classification. However, understanding the demand shift due to ramp meters in an integrated transport system is a key issue to make sense of how the whole transportation system is affected. Previous research has assumed fixed demand when analyzing ramp meters. Ramp metering was first proposed to reduce peak congestion on urban freeways by maximizing throughputs at freeway bottlenecks. For this reason, ramp meters are usually depicted as a short duration (one peak period) demand management device, so its long-term effects on travel demand are, in most cases, overlooked. Moreover, there tends to be a lack of reliable demand data. However, the ramp meter holiday provides us data not previously available to explore the effects of ramp meters on travel demand. Travelers have a full set of possible behaviors in response to ramp metering, summarized in Table 2-2. Changes in departure time, routes and non-work destinations are shorter-term responses. Responses such as relocation are more drastic and hence are not likely to occur due to ramp meters in the short term and also are unlikely to be captured in the 8-week ramp meter holiday. There is no way to analyze those responses via direct freeway traffic counts. Therefore, we focus on the first three types of behaviors.

Table 2-2 Possible Responses to Ramp Metering

Short-term Responses	Long-term Responses
Switch routes	**N**ew Modes
Change destinations (discretionary trips)	**C**ancel Trips (discretionary trips)
Reschedule trips	**A**lter activity sequences
	Change job/house location

Based on 30-sec flow counts collected by loop detectors, three freeway system level statistics can be derived: total trips, total vehicle kilometers traveled and average trip length:

Total Trips: Total trips entering a specific freeway in each 30-sec time interval.

$$Total\ trips = \sum_{n=1}^{N} Q_n \qquad (2\text{-}9)$$

Where:

Q_n flow counts collected by detector n;

N total number of detectors at freeway entrances, including the furthest upstream freeway mainline section and all entrance ramps.

Total Vehicle Kilometers Traveled: VKT on a specific freeway in each time interval.

The basic freeway units used to calculate total VKT are those freeway segments with uniform flow characteristics.

$$Total\ VKT = \sum_{i=1}^{I}\left[\left(\sum_{n=1}^{N_i} Q_n\right) L_i\right] \qquad (2\text{-}10)$$

Where:

I total number of freeway segments;

L_i length of freeway segment i;

N_i number of lanes on freeway segment i;

Q_n flow counts collected by mainline detector n.

Average Trip Length: Average trips length on a freeway in each time interval.

$$Average\ Trip\ Length = \frac{Total \quad Trips}{Total \quad VKT} = \frac{\sum_{n=1}^{N} Q_n}{\sum_{i=1}^{I}\left[\left(\sum_{n=1}^{N_i} Q_n\right) L_i\right]} \tag{2-11}$$

Note that this average trip length is not the actual average OD trip distance because trips starting from a studied freeway may not end on the same freeway and vice versa. So the absolute values of the average trip length can only be used to compare similar cases.

Then those 30-sec total trips, total VKT and average trip lengths are aggregated to seven longer time periods respectively: *(1)* Morning peak: 6:00~9:00; *(2)* Afternoon peak: 15:00~19:00; *(3)* Peak total: *(1)* + *(2)*; *(4)* Weekday off-peak: 00:00~6:00 + 9:00~15:00 + 19:00~24:00; *(5)* Early Morning: 4:00 to 6:00; *(6)* Weekdays; *(7)* Weekend. The median values of the total trips, total VKT and average trip lengths in the above seven durations across the study period are then computed and compared in various ways. The "%Change" results presented in the Result section are percentage changes from values with ramp metering control to values without ramp metering control. By comparing results in different time periods, we are able to answer many key questions about how travel demand responds to ramp metering.

2.2 Measuring Travel Times

This section summarizes the methodology to compute travel times (and speeds) on entrance ramps, freeway segments, and OD pairs on a freeway with and without ramp meters from traffic data collected by loop detectors.

The departure rate (Q_k) arrival rate (q_k) pair in each time interval (k) is obtained from flow detectors for all on-ramps. Throughout the studied peak periods, all ramp arrival detectors have low occupancy readings, indicating no queue spillover effects to local connecting streets. This assures that the delays at on-ramps represent total delays caused by ramp meters. Given arrival rates and departure rates collected at on-ramps, it is possible to find the total travel time every individual vehicle spends at on-ramps by a queuing analysis (Newell 1982), which contains two parts, the free flow travel time from the ramp arrival detector to the departure detector and ramp delay. Since ramps are short, the free flow travel time at ramps can be neglected, and this travel time will be just called ramp delay for the remainder of this chapter (A discussion on distinguishing "delay" and "waiting time in a queue" can be found in Lawson et al. (1997) and Lovell and Windover (1999)).

In the queuing analysis, the ramp delay (d_v) of each vehicle is:

$$d_v = T_v - t_v \tag{2-12}$$

Where:

t_v arrival time of vehicle v;

T_v departure time of vehicle v;

d_v delay of vehicle v (second).

The ramp departure distribution is assumed to be uniform due to the presence of ramp meters. An assumption must be made about the arrival distribution of vehicles at the back of the queue. A uniform arrival rate would give a lower bound for delay. A more reasonable assumption is to use a Poisson arrival process, which allows for bunching of vehicles. If the number of vehicles arriving at a queuing system has a Poisson distribution with a mean of q_k customers per unit of time, the time between arrivals has a negative exponential distribution with a mean of $1/q_k$. With the data collected by departure and arrival detectors, the ramp delay for each vehicle and the average ramp delay for vehicles arriving in each 5-minute interval can be obtained. Headways are simulated using a negative exponential distribution for 50 times. Ramp delays are computed using equation (2-12) and then averaged over all simulation runs.

Freeway mainline loop detectors provide the flow and occupancy information in aggregated 5-min intervals (raw data are in 30-sec intervals, see section 5), useful for computing traffic flow on freeway segments. Based on these data, the space-mean speeds in every time interval can be computed as in equation (2-13):

$$u_{m,k} = 2.76\times10^{-5}\left(\frac{Q_{m,k}l_m}{O_{m,k}}\right) \tag{2-13}$$

Where:

m index of detectors at a freeway detection station, one detector per lane;

$u_{m,k}$ space mean speed at the m^{th} detector in time interval k (km/sec);

$Q_{m,k}$ flow of the m^{th} detector in time interval k (vehicles/hour);

l_m average vehicle length plus the length of the m^{th} loop detector (m);

$O_{m,k}$ time occupancy of the m^{th} detector in time interval k (%).

Average vehicle length plus detector length, commonly known as the effective vehicle length is a crucial factor in estimating speed from inductive loop detector flow and occupancy readings. Average vehicle length estimates were taken from the Mn/DOT TMC effective loop detector length calibration/normalization study. Because freeways have multiple lanes, there is more than one loop detector at each station, and there are multiple space-mean speeds (one for each lane) at each station derived from equation (2-13). The weighted mean of all lanes will be used as the speed at each station, which is:

$$\bar{u}_{i,k} = \frac{\sum_{m=1}^{M_i} K_{m,k} u_{m,k}}{\sum_{m=1}^{M_i} K_{m,k}} \qquad (2\text{-}14)$$

Where:

i index of freeway detection stations, one station per freeway segment;

M_i number of detectors at station i;

$\bar{u}_{i,k}$ weighted space mean speed at station s in time interval k;

$K_{m,k}$ density of the m^{th} detector in time interval k.

In theory, weighting by density produces space mean speed. It is noted in our study that if averaging by flow the mean speed results are almost exactly the same as those weighting by density (with a difference less than 1 km/h), even when freeways are congested. Assume that this speed is the average speed within this segment (that is, speeds within a segment are uniform). If there are no detectors in a segment, the speed at the nearest station will be taken to be the average speed in this segment (though this kind of situation rarely happens). Then freeway segment travel times in each time interval can be obtained:

$$t_{i,k} = L_i / \bar{u}_{i,k} \tag{2-15}$$

Where:

$t_{i,k}$ travel time on freeway segment *i* in time interval *k* (second);

L_i length of freeway segment *i* (km).

Once the average delay at ramps and the average travel time on freeway segments in each time interval are determined, it is possible to build an OD trip travel time matrix for a freeway system during a whole peak period (origins: starting points of on-ramps; destinations: off-ramps). However, this requires that the previously calculated ramp delays and freeway segment travel times be synchronized. First a database which records ramp delays and freeway segment travel times in each 5-minute interval is built. Then trip travel times for each OD pair (from on-ramp *r* on freeway segment *i* to an off-ramp on freeway segment *j*) are calculated. The OD trip travel times are calculated by equation (2-16):

$$t_{i,j,k} = d_{r,k} + t_{i,k+x_1} + t_{i+1,k+x_2} + \ldots + t_{j-1,k+x_{j-i}} \tag{2-16}$$

Where:

$t_{i,j,k}$ travel time from origin *i* to destination *j* departing in time interval *k*;

$d_{r,k}$ delay on ramp *r* in time interval *k*;

$t_{i,k+x1}$ travel time on freeway segment *i* in time interval ($k+x_1$);

$x_1, x_2, \ldots, x_{j-i}$: synchronization coefficient, *x* is equal to the integer part of the result of the travel time from the origin to the beginning of this segment divided by the time interval (300 seconds).

Provided OD distances, an OD trip speed matrix can be derived from the OD trip travel time matrix.

These OD trip mobility matrices, along with ramp delays at all on-ramps, freeway segment travel times and ramp/freeway segment flow data, are then used to compute various performance measures introduced in the previous section for both the metering-on and the metering-off cases.

2.3 Data

The raw data are archived 30-sec flows and occupancies provided by the Metro Division Traffic Management Center of the Minnesota Department of Transportation. All data used in this study have passed TMC's continuity test for detector readings, which is an algorithm based on flow conservation to check the accuracy of detector data. In choosing the study locations and the study period, we follow these criteria:

(1) The time/space domain must be free of congestion at the beginning and the end of the study period. Thus the flow conditions within selected freeway sections during the study period will not be affected by further downstream queues.

(2) Data must be complete. There are no malfunctioning detectors in the selected freeway sections during the study period.

Under criteria (1) and (2), the study period was chosen to be the afternoon peak from 14:30 to 19:30 and only four freeways are qualified: I-494 outer loop (eastbound and southbound), I-494 inner loop (westbound and northbound), TH169 northbound and TH62 westbound (see Figure 2-3). Many other freeways did not qualify because there were no working on-ramp queue detectors which could provide arrival data in computing ramp delays for the metering-on case. Even on the selected four freeways, there were still some on-ramps where queue detectors either over-count or under-count vehicles. These on-ramps are excluded from the analysis.

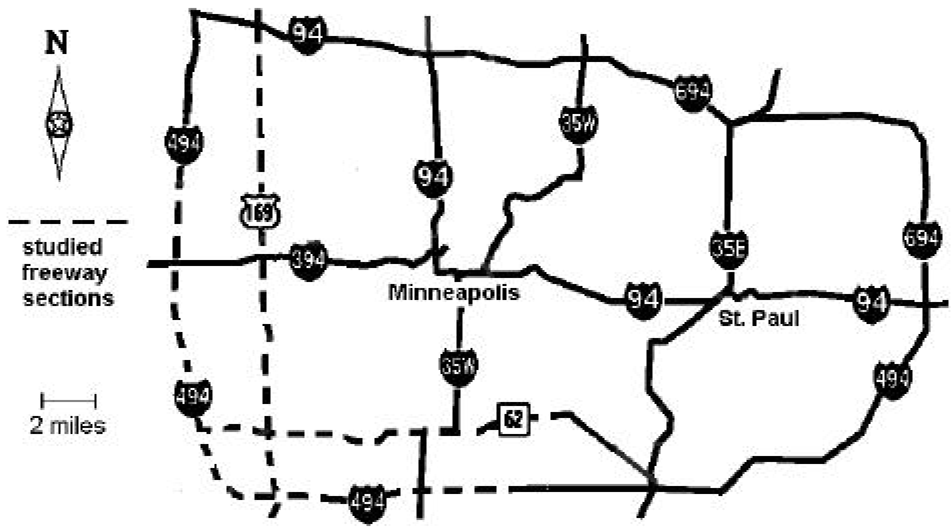


Figure 2-3 Study Locations

(3) Only use the data collected during the last six weeks of the holiday to avoid the noisy and unrepresentative travel behaviors immediately after shut-off. Data collected in the corresponding six weeks (from the fourth week in October to the first week in December) in year 1999 are used to compute performance measures for the metering-on case in order to avoid impacts of seasonal demand fluctuation.

(4) “Bad-weather” days were excluded from the two six-week periods (metering-on and metering-off). Days with more than 1 centimeter of hourly precipitation (rain or snow) in any 1-hour duration in the studied afternoon peak period are “bad-weather” days. For example, if the Tuesday in the first week of October 2000 is identified to be a “bad-weather” day, this Tuesday and the corresponding Tuesday in October 1999 will both be excluded from the analysis, no matter the weather condition on the 1999 Tuesday. We only considered precipitation since freeway capacities in the Twin Cities were found to be most sensitive to the precipitation among all weather factors (wind, temperature, sky cover, etc) (Ries 1980). Weather data were from NOAA hourly weather reports collected at Minneapolis and St. Paul Airport Station. While it would be interesting to examine bad weather conditions, we are unable to separate seasonal effects, weather effects and control effects.

After the above data-filtering process, we now have a data set collected on 29 "good-weather" days out of six weeks for both metering-on and metering-off cases on selected freeways (on TH62, 27 days because two more days are excluded due to malfunctioning detectors). Various performance measures are computed based on this data set.

2.4 Evaluation Results

2.4.1 Mobility, Equity, Consumers' Surplus, Productivity and Accessibility

The computed mobility (travel time, speed), equity, consumers' surplus, productivity and accessibility measures on the four selected freeways are very similar. Therefore, in the following presentation, we will only discuss in detail the results obtained on a representative freeway, TH169. Important statistics are tabulated for all four studied freeways.

First the relationship between mobility and equity with and without ramp meters is estimated. Figures 2-4 and 2-5 illustrate the trends in the change of mobility and equity with and without metering on TH169. The mobility and equity values in the graph are averaged mobility and equity measures over all 29 days. For instance, to obtain spatial speed/Gini coefficient for time interval k in Figure 2-4, we compute speeds/Gini coefficients in time interval k for all 29 days. It is the average of these speeds/Gini coefficients that is presented in the graph. Figure 2-4 looks at spatial equity vs. spatial mobility, and shows worsening equity and mobility as the peak is reached. More interesting, the tradeoff between mobility and spatial equity is evident. That is the metering-on case has better mobility but worse spatial equity compared to the metering-off case.

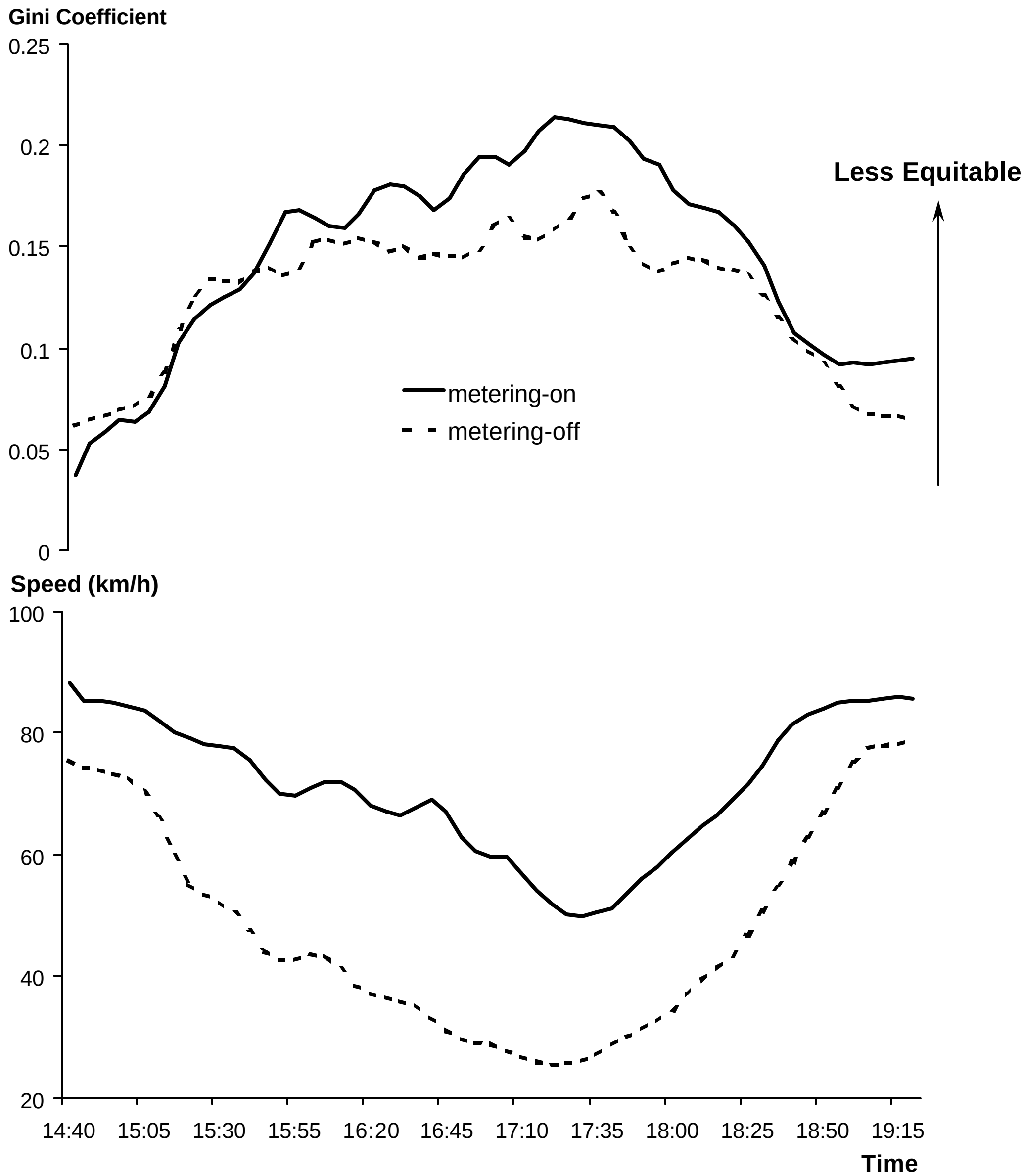


Figure 2-4 Spatial Equity and Mobility (Speed) on TH169

In Figure 2-5, temporal equity vs. temporal mobility, some short trips (those on the left side of the graph) actually are hurt in both mobility and temporal equity by ramp metering, while the longest trips (those on the right side) benefit the most. Thus it is fair to say that drivers making shorter trips are losers and those making longer trips are winners under the ramp metering control policy.

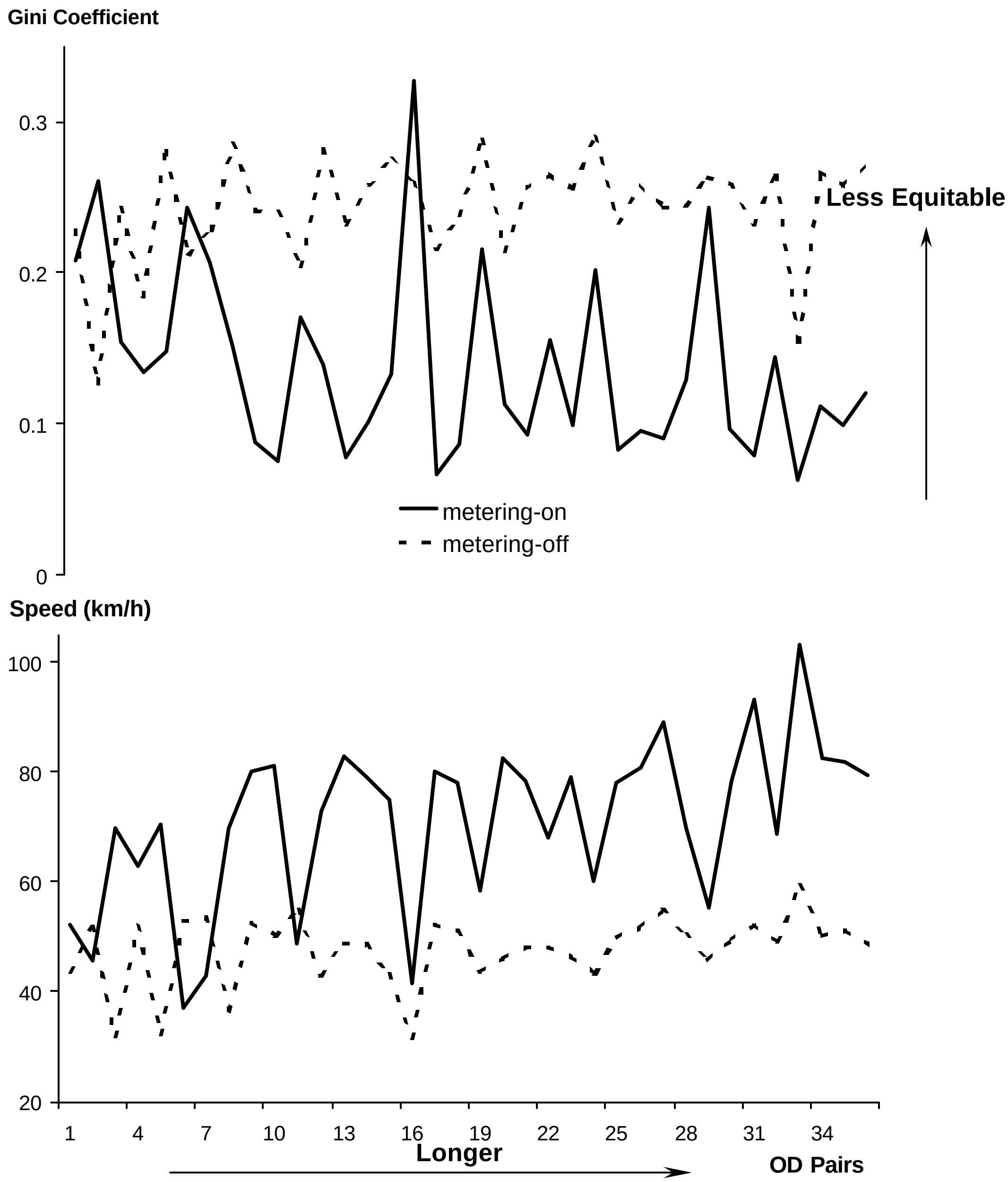


Figure 2-5 Temporal Equity and Mobility (Speed) on TH169

Table 2-3 Spatial-Temporal Equity and Mobility

Freeways		Travel Speed (km/h)		Travel Time (sec/km)	
		Metering-Off	Metering-On	Metering-Off	Metering-On
TH169NB	Mobility	48	71	104	66
	Gini	0.269	0.188	0.285	0.284
I494NW	Mobility	65	78	63	57
	Gini	0.156	0.170	0.178	0.255
I494SE	Mobility	53	75	97	62
	Gini	0.241	0.141	0.285	0.239
TH62WB	Mobility	75	83	51	50
	Gini	0.097	0.127	0.112	0.205

What ramp metering delivers in terms of mobility and equity for all OD pairs throughout the afternoon peak period can be shown by the comparison of spatial-temporal mobility and equity of the two cases in Table 2-3. The spatial-temporal mobility measures presented in Table 2-3 are essentially the same as overall system performance measures which appear frequently in other freeway performance studies. Previous studies indicate that ramp meters can increase the mobility of freeway networks (Taylor and Meldrum 2000, Haj-Salem and Papageorgiou 1995, among others), which is confirmed here. With ramp metering, average travel time per kilometer of travel (taking ramp delay into account) on TH169 decreases from 104 sec to 66 sec while average travel speed increases from 48 to 71 km/hour. The spatial-temporal equity results are mixed, favoring neither the metering-on case nor the metering-off case. This can be explained by the results presented in the previous two paragraphs: (1) spatial equity decreases with ramp metering; (2) temporal equity of short trips decreases with ramp metering; (3) temporal equity of long trips increases with ramp metering.

Daily productivity, changes in consumers' surplus and accessibility are derived first using the method introduced in section 2. Then the average daily productivity, changes in consumers' surplus and accessibility over all 29 day are computed and presented in Table 2-4, 2-5 and 2-6 respectively.

Table 2-4 Productivity

Freeways	Productivity (km/hour)		
	Metering-Off	Metering-On	%Change
TH169NB	46	71	53
I494NW	67	83	22
I494SE	54	78	41
TH62WB	80	90	12

Table 2-5 Changes in Consumers' Surplus

Freeways	Average Daily Changes in Consumers' Surplus (vehicle*hours)		
	On-Ramps	Freeway Mainline	System
TH169NB	-237	2534	2298
I494NW	-189	1551	1363
I494SE	-86	2535	2445
TH62WB	-20	348	328

Table 2-6 Accessibility

Freeways	Model 1: $f(C_{ij}) = 1/t_{ij}^2$			Model 2: $f(C_{ij}) = e^{-0.00189 t_{ij}}$			Model 3: $f(C_{ij}) = e^{-0.08 t_{ij}}$		
	Off	On	%	Off	On	%	Off	On	%
TH169NB	1064	2073	94	47425	47125	0	21111	27521	30
I494NW	2064	3127	52	35509	37733	6	17997	19895	11
I494SE	1540	2865	86	38225	39969	5	16607	21714	31
TH62WB	2554	3262	28	45896	46666	2	30644	32147	5

On: with ramp metering; *Off:* without ramp metering; *%:* percentage changes

Table 2-4 shows productivity, the vehicle kilometers of travel per vehicle hour of travel on all four studied freeways. Productivity on freeway mainline increases with ramp metering but productivity at on-ramps decreases. Combining freeway mainline with on-ramps gives a system productivity measure ((mainline *VKT* + ramp *VKT*)/(mainline *VHT*

+ ramp *VHT*)). The system productivity is improved immensely by ramp metering. The percentage increase on TH169 is 53%.

Table 2-5 summarizes changes in consumers' surplus on all four studied freeways. The changes in consumers' surplus for each individual freeway mainline segment on TH169 are summed to get a benefit from metering of 2534 vehicle hours. The change in consumers' surplus on entrance ramps is found to be -237 vehicle hours. As expected, ramp meters significantly reduce the consumers' surplus of entrance ramps. However, ramp meters benefit the freeway mainline more than they harm entrance ramps, so an overall positive change in consumers' surplus of 2298 vehicle hours is recorded.

Three different accessibility models were applied to TH169. The first is a classic gravity model, the second a model estimated for freeways in the Twin Cities, and the third from a regional gravity model estimated by the primary author for Washington DC (Levinson and Kumar 1995). For all the three models, accessibility increases with ramp metering, as shown in Table 2-6.

2.4.2 Travel Time Variation

Inter-Day travel time variation results for four studied freeways are shown graphically in Figure 2-6 (range/median plots). It is obvious that for most OD pairs (103/127: 26/45 OD pairs = 5 km, 77/82 OD pairs > 5 km), inter-day travel time variability is reduced by implementation of the ramp metering system (t tests of $V_{off} - V_{on} > 0$ are statistically significant at level 0.01). Freeway peak hour travel reliability increases. However, for extremely short trips (= 5 km), it is hard to say whether ramp meters improve or reduce inter-day travel time variations. Figure 2-7 illustrates intra-day travel time variation results with two curves representing the metering-on and metering-off cases respectively. Although the intra-day travel time variation reductions caused by ramp meters differ by freeway segment, clearly ramp meters play a positive role in reducing intra-day travel time variation. As in the inter-day results, intra-day travel time variation of long trips is reduced more significantly than those of short trips.

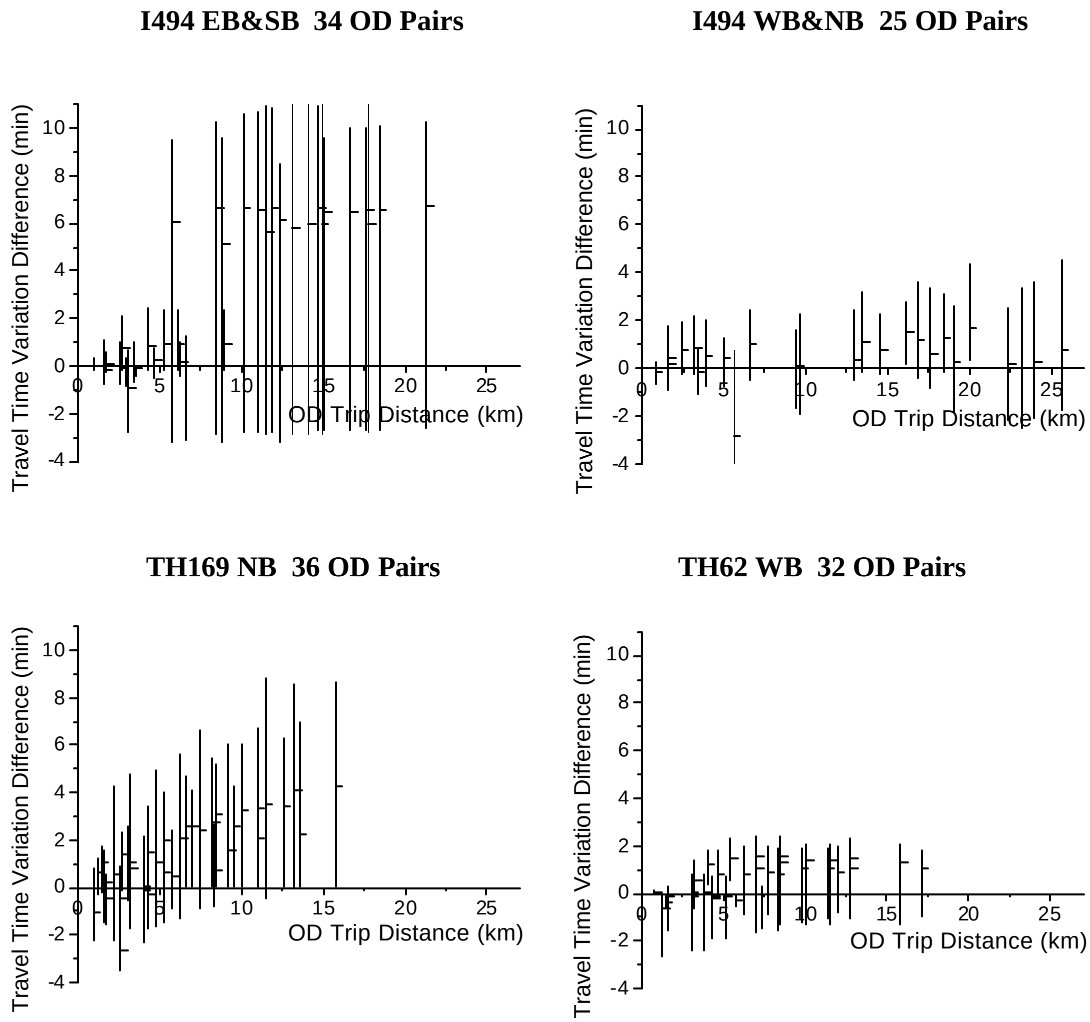


Figure 2-6 Inter-Day Travel Time Variation Difference of Each Freeway

Without Ramp Meters – With Ramp Meters ($V_{off} - V_{on}$)

During the ramp meter holiday, only variation of freeway mainline conditions cause variability in travel times. But for a metered freeway, ramp delay variability also contributes to the overall trip travel time variation. The contribution of ramp delay variability to the overall travel time variation can be separated out. Results show that for a significant number of trips (42/131 OD Pairs), inter-day travel time variation of these trips could be reduced more than 50% with a metering strategy that ensured the same ramp delay at the same time-of-day, across days. In other words, on different days, the Minnesota algorithm results in quite different ramp delays even during the same time intervals.

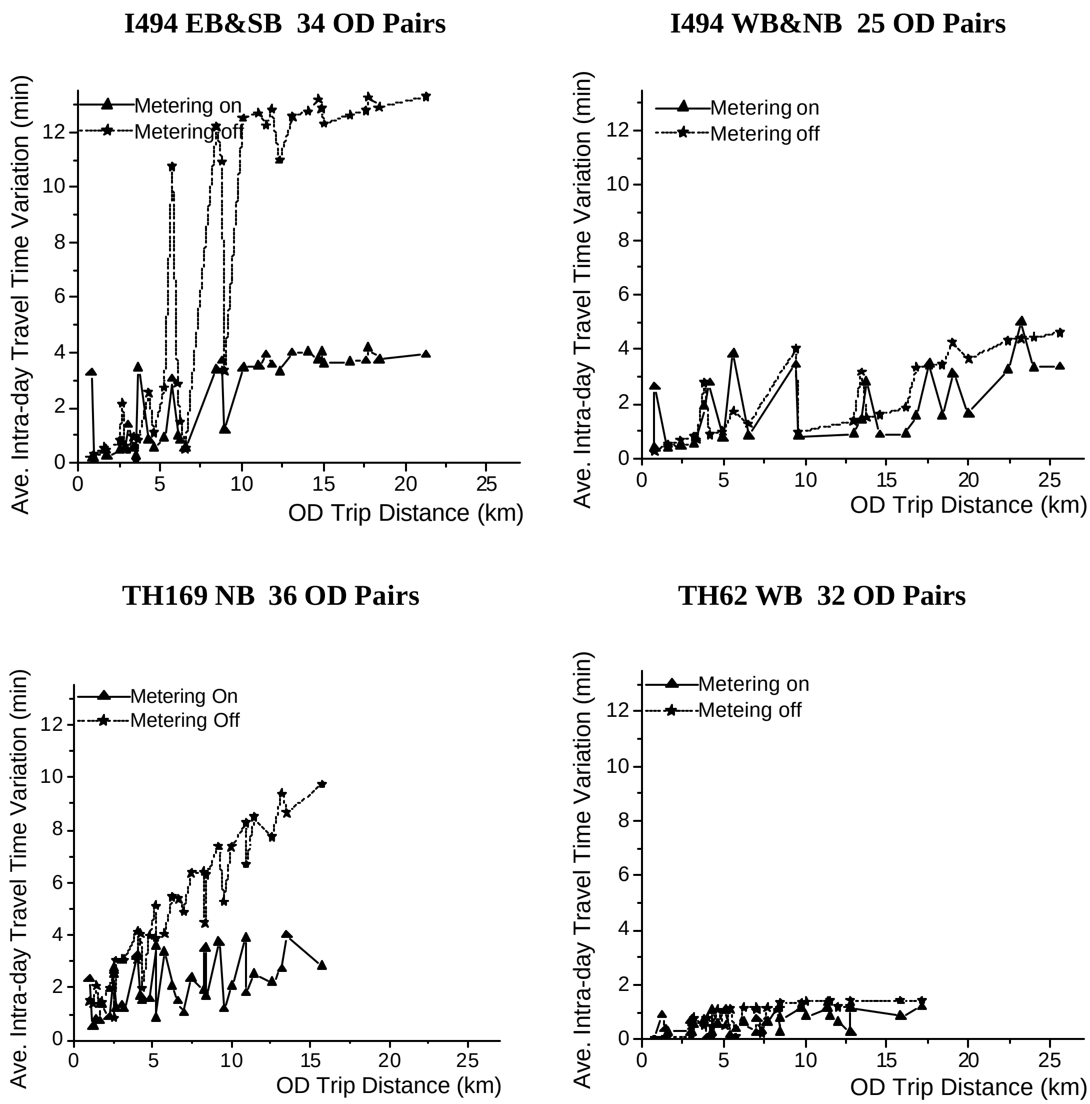


Figure 2-7 Intra-Day Travel Time Variation: with/without Ramp Metering Control

Assuming that all OD pairs have the same number of trips, the difference in overall average inter-day travel time variation is 1.82 minutes. That means the implementation of ramp metering control can reduce 1.82 minutes of travel time standard deviation. Black and Towriss (1993), Small (1995) and Small et al. (1999) estimated a so-called “reliability ratio” (ratio of cost of standard deviation to mean travel time when scheduling costs are not separately considered) and consensus results of 0.7, 1.27 and 1.31 were obtained. Using the unit cost of travel time uncertainty — $0.21 per minute of standard deviation which was estimated by Small et al. (1999) from a stated preference passenger survey in Los Angeles, ramp metering can save $0.38 for each freeway trip in terms of increased travel certainty.

2.4.3 Demand Responses

Because the ramp meter shutdown experiment was carried out in 2000 and the chosen corresponding metering-on period is in 1999, our results will be drivers' response to metering holiday. There is no reason to believe that the results are not reversible and it is expected that travelers' response to the return of ramp meters would simply be the opposite.

Spreading of the Peak

The ratio of the peak 15-minute flow to the peak hour flow is widely used to represent the spreading of the peak. This can also be done using loop detector counts by aggregate traffic flows at 15-minute time intervals. However, this is only one way to look at peak spreading and some other methods are also useful. Here, we make two comparisons. First, we compare the change of total traffic flows (total trips) in the early morning period and those in the morning peak. It is obvious from Figure 2-8a that travelers respond to the shut-off of ramp meters by departing earlier in the morning to avoid congestion in the morning peak periods. Secondly, we compare the total vehicle kilometers traveled in two peak periods with those in two off-peak periods in Figure 2-8b: weekday off-peak and weekends. When ramps are unmetered, relatively more travel is made during off-peak hours while less travel is made in the peak period. Many trips that occur in the peak periods with metering are pushed entirely to the off-peak periods without metering, generally weekday off-peak hours. The small increase in VKT on weekends is in tune with typical traffic growth.

Figure 2-8a Early Morning and Morning Peak Percentage Change of Total Trips

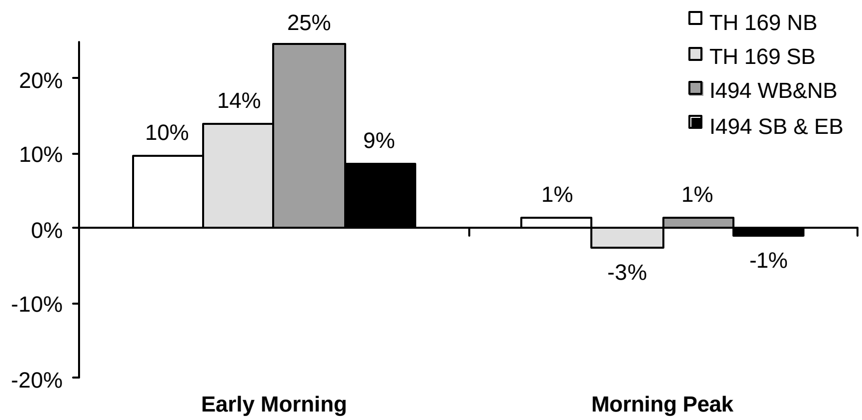
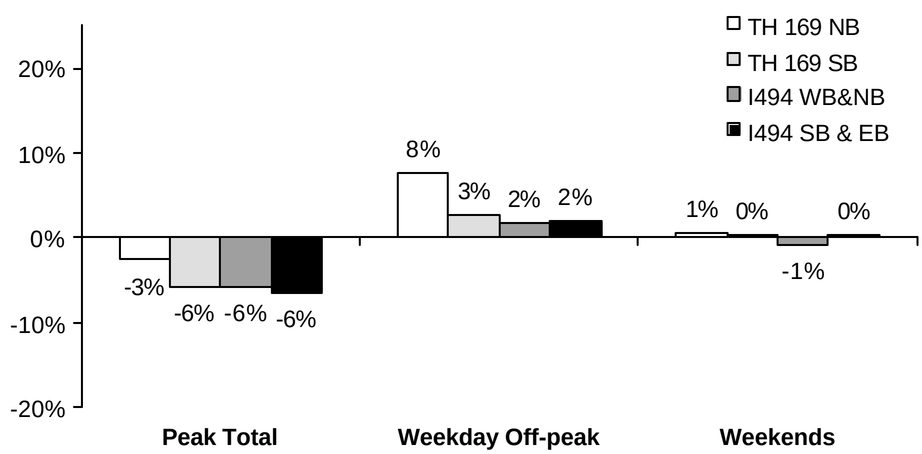


Figure 2-8b Peak Total and Weekday Off-peak Percentage Change in Total VKT



Non-work Trips

While work trips are characterized by a strongly desired arrival time and a fixed destination, non-work trips can more easily shift time and location. Research comparing the United States Nationwide Personal Transportation Survey reported a significant amount of discretionary travel (Pisarski 1992). Therefore, it is rather important to explore how non-work trips react to the shut-off of ramp meters. Also given that previous mobility results show different effects of ramp meters on short trips and long trips, it is also of high interest to separate out short non-work trips and long work trips to see their different demand responses.

A direct traffic count can reflect trip purpose only if it is collected at certain locations, such as a street to a recreation center. However, some inferences can be made based on the assumption that morning peak trips are dominated by work trips but afternoon peak trips consist of both work trips and non-work trips. Non-work trips, such as shopping and visiting are more likely to happen in the afternoon than in the morning. Analysis in Montgomery County, MD also confirms this point (Levinson and Kumar 1994): about 80 percent of travel during the morning peak is classified as work trips; while in the afternoon peak period, work travel constitutes only 50 percent of all trips. Then the difference between the morning peak and the afternoon peak in terms of total trips and total vehicle kilometers traveled can be ascribed to a large extent to non-work trips that occur in the afternoon peak. For instance, the afternoon peak non-work trips on a northbound freeway will be the difference between the total afternoon peak northbound trips and the total morning peak southbound trips, because work trips reverse directions between morning and afternoon.

As expected, afternoon peak total trips and total VKT are always greater than those in the morning peak period, largely explained by the additional non-work travel during the afternoon peak period. There are more non-work trips in the afternoon peak period as a result of shutting off ramp meters (see Figure 2-8c). The increase in afternoon non-work trips cannot be explained by the weekday trip growth rate, the four bars from the right end in Figure 2-8c. It is even more interesting to look at the change in total VKT. It decreases despite the increase in total trips. The only explanation is that as a result of the holiday, long peak-period non-work trips are discouraged while short non-work trips are encouraged. On average, during the holiday people are making more but shorter non-work trips in the peaks, exhibited by the decrease of average non-work trip length. Although this effect for all afternoon trips as a whole is mitigated to some extent by the inelasticity of work-trips, this trend of increasing total trips but decreasing total VKT still can be seen for the whole afternoon peak-period.

Figure 2-8c Afternoon Peak Non-work Trip Percentage Changes in Total Trips, Total VKT and Average Trip Length

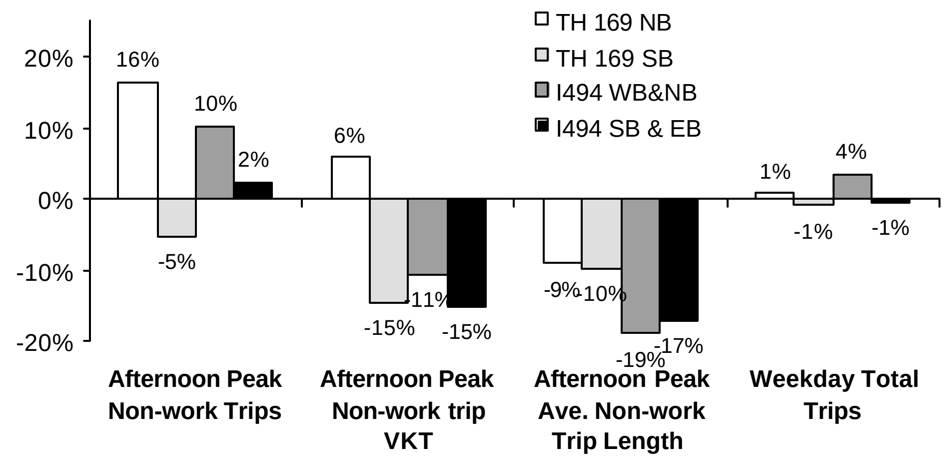
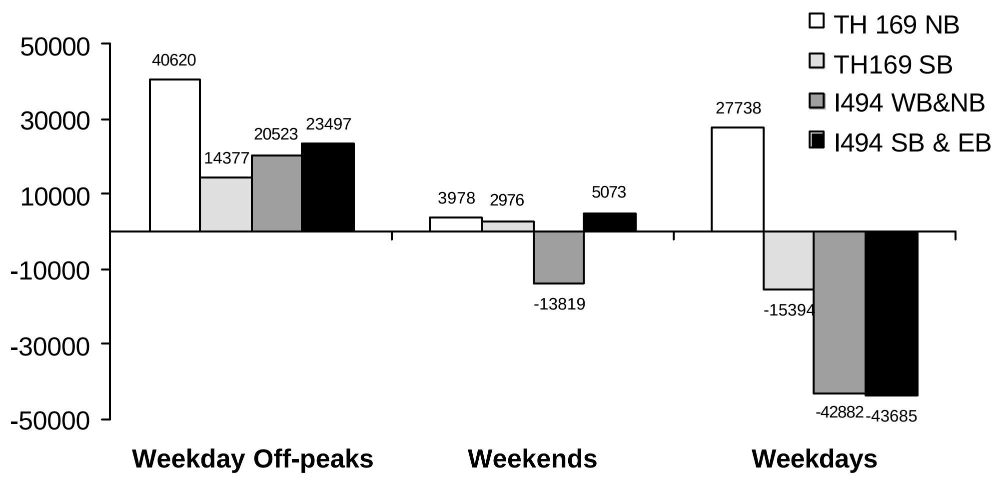


Figure 2-8d Changes in Total VKT of Weekday Off-peaks, Weekends and Weekdays



Travelers reschedule their short non-work trips to the peak period and defer their long non-work trips to off-peak times when the ramp meters are off. There is no evidence that travelers change their weekend short non-work trips to weekday peaks. So, the additional short non-work trips in the afternoon peak without metering should occur during the weekday off-peak with metering. However, things are not so straightforward when we think about when those long non-work trips are rescheduled. They can then occur in weekday off-peak hours or weekends instead. Figure 2-8d shows the changes in total VKT for these two time periods. Considering that the loss of short non-work trips should decrease the total VKT in the weekday off-peaks and that the increase of total VKT resulting from peak spreading only constitutes a small part of the total increase of VKT in the weekday off-peaks, we draw the conclusion that many long non-work trips are redistributed to weekday off-peak hours or vanish entirely.

Changing routes and destinations

Figure 2-8b shows that freeways carry less traffic in peak periods without metering and people reschedule their trips to weekday off-peak periods. If all trips pushed out from peak period are rescheduled to occur during off-peak hours, the total change in VKT with/without ramp meters should be non-negative taking traffic growth into account. However, if we take a look at the change in total weekday travel (see Figure 2-8d), the total weekday VKT decreases on most freeways. It is evident that travelers react to congestion without metering not only by rescheduling their trips but also by changing routes and destinations. They may switch to a local arterial, a less congested freeway or even a less congested direction. Our results cannot provide more details. In general, unmetered freeways take less travel, but not much less.

2.5 Implications on Developing New Ramp Control Strategies

This chapter analyzed ramp metering in the Twin Cities using a number of measures of effectiveness. In general, the findings were favorable to ramp metering on four studied freeways.

Prior to October 2000, the Minnesota ramp meter control strategy had been focused on the freeway mainline, ignoring delay on ramps. This theory of freeway ramp metering argues that ensuring higher flows on freeways guarantees lower flows elsewhere on the road network. The more vehicles able to use the freeways, the fewer that will burden local streets. By the objectives the ramp meters were intended to satisfy, maximizing freeway throughput at selected bottlenecks, this study confirms that the facility performs better in the presence of operating ramp meters than in their absence. The change in consumers' surplus was positive while at the same time, the productivity of the system almost doubled on some freeways. Travel speeds (after considering ramp delays) and flows are consistently higher with ramp metering than without. While this is logical if the intention is to maximize freeway throughput, it does not necessarily maximize user satisfaction of the system, especially when the ramp delays are far from being evenly distributed. For instance, spatial equity is worse with ramp metering.

Ramp meters are particularly helpful for long trips relative to short trips. This study on TH169 and three other freeways shows that trips more than 3 exits in length benefit, while many trips 3 exits or less are hurt by ramp meters. Overall, longer trips benefit more than shorter trips. Originally, freeways (the *interstate* highway system) were designed to serve intercity not local traffic. However freeways have evolved to serve commuters. Whether reserving the freeways for long trips is still an appropriate objective is an important public policy debate.

Other possible objectives or constraints, such as ensuring that delay is distributed as evenly as possible, are not achieved by the ramp metering control in place before this study. Consider for instance traffic signals on streets. If a minor road meets a major road, and the major road is operating at capacity, it might be most efficient (in terms of minimal total delay) to give 100% of the green time to the major road and 0% to the minor road. This is not done. Traffic signals alternate back and forth in part to ensure some equity, so that travelers on the minor road do not have an excessive wait. Generally cycle lengths on roads at capacity are under 3 minutes, though this varies. A similar limit on individual delay, even at the expense of overall freeway efficiency, may be necessary for ramp meters to satisfy equity considerations. Future research should pursue alternative strategies. Furthermore, if efficiency is defined more broadly, for instance, maximizing the utility of travelers, recognizing a non-linear value of time, that constraint may be even more essential. So while the Minnesota algorithm may in generally be better than no metering at all, there may be other control strategies that perform better still on a wide variety of performance measures.

The author believes that a ramp metering system that satisfies users must consider ramp delay in addition to freeway throughput. Unfortunately current data does not permit any such strategies. Aside from the few locations used in this study, there exists no accurate measure of the number of vehicles waiting in queue at a ramp meter at any given time in the Twin Cities.

Ramp metering was designed to improve freeway traffic flow and safety. While it generally does both (see Cambridge Systematics (2001) for a discussion of safety effects), it also has the affect of improving travel time reliability for long trips. This should be captured in the analysis of ramp metering benefits.

Travelers react to the congestion resulting from the ramp metering shut-off by departing earlier in the morning and rescheduling their trips to weekday off-peaks, which shows a certain amount of peak spreading. Travelers also change routes and destinations. In general, freeways handle fewer vehicle kilometers when there is no ramp metering, but not much fewer. A unique demand response to ramp metering appears on non-work trips. Although there are more non-work trips during the afternoon peak period in general, total afternoon peak non-work trip VKT decreases when ramp meters are shut off. The absence of ramp meters discourages long peak-period non-work trips, which are deferred to weekday off-peak hours. On the other hand, short non-work trips are encouraged to occur in the peak period without meters. On average, people make more but shorter non-work trips in the afternoon peak periods on unmetered freeways. This is because long distance trips save time at the expense of short trips when freeways are metered as mentioned before.

In summary, the above discussion has the following implications on developing new ramp metering control strategies:

- The theory of ramp metering in the Twin Cities argues that ensuring higher flows on freeway mainlines guarantees lower flows elsewhere on the road network. While this is logical if the intention is to maximize freeway throughput, it does not necessarily maximize user satisfaction of the system, especially when the ramp delays are not evenly distributed.
- In an attempt to evenly distribute ramp delays, the Minnesota algorithm uses historical data to estimate demands on individual ramps. The evaluation results suggest that it is not an effective way to equalize ramp delays. Traffic demand fluctuations, especially those within a peak period, cannot be captured well by

historical data. Therefore, some real-time data collection efforts are necessary to obtain accurate enough queuing information at on-ramps. Such data collection work can be accomplished by either departure/arrival detector pairs or video capture and processing techniques.

- The trade-off between efficiency (measured by travel time) and spatial equity is evident. A ramp control strategy minimizing total system travel time may not be acceptable to the public due to its poor equity. There are two potential ways to solve this problem:

  (1) Keep the metering objective as minimizing total travel time. In addition, add a maximum-ramp-delay constraint.

  This approach considers equity in a constraint equation, which may bring some difficulties to the solution process of the whole optimization program in theoretical studies. However, if this constraint is left out of the solution process and rather serves as an overriding rule over the solution from particular optimization programs, the implementation of it could be very straightforward.

  (2) Use other metering objectives. For instance, the objective function could minimize total non-linearly weighted travel time so that longer delays have higher weights than shorter delays which have higher weights than free-flow travel times.

  With equity considerations built into the metering objective function via a non-linear weighting process, this approach is theoretically preferable. Our current knowledge on non-linear travel time, however, is not sufficient. Future studies exploring appropriate value-of-time functions for travel delays are merited.

- Ramp metering also has the effect of improving travel time reliability. Although current methods for evaluating travel time reliability are not conclusive, this benefit of ramp metering should be kept in mind in developing ramp control strategies. We have not yet found effective ways to further improve travel reliability through ramp metering, but a ramp control strategy that cannot significantly improve travel time reliability over the no-control scenario should be considered undesirable.

- The evaluation results also confirm some hypotheses on the demand responses to ramp metering. Commuters make more but shorter non-work trips in the afternoon peak periods on unmetered freeways. Travelers also change routes and destinations. These demand shifts should benefit the whole transportation system. One demand response, which is not previously hypothesized and probably counteracts the effectiveness of ramp metering, is that the absence of ramp meters results in peak spreading.

3 FORMULATING OPTIMAL RAMP CONTROL PROBLEM

In this chapter, a new formulation of the time-dependent optimal ramp control problem based on linear programming is developed. The virtue of this new formulation is that it dose not require any OD information. All input variables for the new linear programming formulation are directly measurable by inductive loop detectors in real-time. A heuristic solution to the rolling-synchronized-horizon version of the new formulation is developed which is non-anticipatory (it only uses information that is known at each decision point), efficient and shown to be the global optimal solution on a small-scale network when certain assumptions hold. The heuristic solution suggests that, when efficiency is measured by the total travel time in the freeway system i.e. freeway mainline and ramps, the most efficient ramp control strategy is also the least equitable strategy.

The heuristic solution also enables us to build an analytical framework for ramp metering studies under which different research results could become more comparable. This analytical framework will be introduced in Chapter 4. The optimal ramp control theory that will be presented in this chapter is also the theoretical foundation for the development of efficient and equitable ramp control strategies in later chapters.

3.1 Notation

We will follow these notation conventions in explaining the optimal ramp control problem and in discussing the new ramp metering algorithms:

A:* the furthest upstream flow of the freeway system;

B: flow on a freeway section measured at the furthest downstream point;

B:* the furthest downstream flow of the system[2].

C: capacity of a freeway section;

D: arrival flow at an entrance ramp;

---

[2] In fact, $A^* = M_I + B_K$, and $B^* = B_1$. We still assign separate notation for them to simply some equations.

i: index of entrance ramps;

I: number of entrance ramps;

j: index of exiting ramps;

J: number of exiting ramps;

k: index of freeway sections;

K: number of freeway sections;

M: departure flow at an entrance ramp (also the metered flow if the ramp is metered);

q: arrival rate for the whole freeway system;

Q: departure rate of the whole freeway system

S: on-ramp demand (sum of the standing queue in the previous time interval and the arrival rate in the current time interval, see equation (3-9) and (3-10));

t: index of time intervals;

t_0: starting time of the ramp metering control period;

T: end time of the control period;

X: flow at exiting ramps;

a: diversion rate of an exiting ramp = exiting ramp flow / upstream mainline flow, $\alpha \in [0, 1]$;

g: a $I \times K$ indication matrix, $\gamma_{ik} = 1$ if entrance ramp i is on section k,

$\gamma_{jk} = 0$ otherwise;.

d: a $J \times K$ indication matrix, $\delta_{jk} = 1$ if exiting ramp j is on section k,

$\delta_{jk} = 0$ otherwise;

D: a $K \times K$ synchronization matrix, $\Delta_{k1,k2}$ = free-flow travel time from section k_1 to k_2;

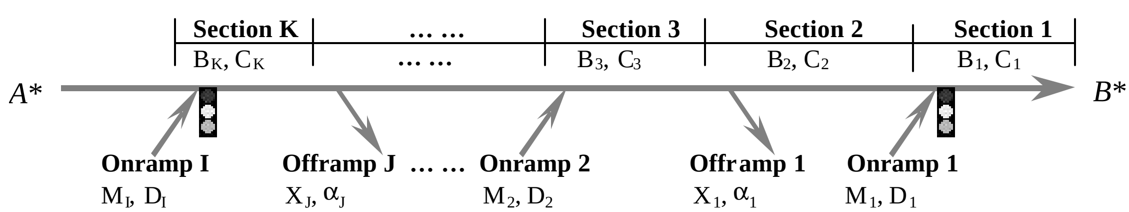


Figure 3-1 Coding a Typical Freeway System

Figure 3-1 illustrates the freeway coding conventions in this report. A freeway system can be divided into a series of freeway sections (section 1 to section K). Each section contains either an on-ramp or an off-ramp. On-ramps could be metered by a particular control algorithm or not metered. A section starts from the location immediately upstream of the on/off-ramp it contains and ends at the starting point of the immediately downstream section. All sections can be categorized into three types: metered-section (e.g. section 1 and section K), unmetered-section (e.g. section 3) and off-ramp-section (e.g. section 2).

3.2 Formulating the Optimal Ramp Control Problem Based on Linear Programming

We consider the whole freeway system as a queuing system (see Figure 3-2). The arrival rate of this queuing system at any time interval t (q_t) is:

$$q_t = A_t^* + \sum_{i=1}^{I} D_{it} \quad (3\text{-}1)$$

The departure rate of this queuing system at any time interval t (Q_t) is:

$$Q_t = B_t^* + \sum_{j=1}^{J} X_{jt} \quad (3\text{-}2)$$

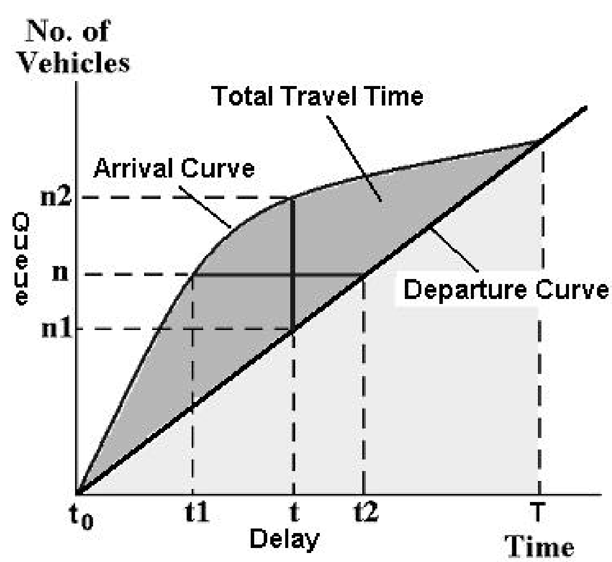


Figure 3-2 Freeway System Queuing Diagram

The total absolute travel time in this queuing system is the area bounded by the arrival curve and the departure curve in figure 3-2[3]. Our objective is simply to minimize the total travel time (heavily shaded area). By assuming a fixed arrival curve, i.e. no diversions or route choices, minimizing total travel time is equal to maximizing the lightly shaded area under the departure curve, which is the integral of departure rate over the whole control period:

$$Max \int_{t=t_0}^{T} Q_t dt \tag{3-3}$$

Papageorgiou (1983) shows that the time-discrete version of this integral is just:

$$Max \sum_{t=t_0}^{T} [(T-t) \cdot Q_t] \tag{3-4}$$

Substitute equation (3-2) into (3-4), the objective function then becomes:

$$\underset{M}{Max} \sum_{t=t_0}^{T} [(T-t)(B_t + \sum_{j=1}^{J} X_{jt})] \tag{3-5}$$

This objective is subject to a set of constraints:

$$B_{k,t} < C_{k,t} \quad \textit{for all k and all t} \tag{3-6}$$

$$0 \le M_{i,t} \le S_{i,t} \quad \textit{for all i and all t} \tag{3-7}$$

We also have the following relationships:

$$X_{jt} = \sum_{k=1}^{K} (d_{jk} B_{kt} \frac{a_{jt}}{1-a_{jt}}) \tag{3-8}$$

$$S_{i,t} = S_{i,t-1} + D_{i,t} - M_{i,t-1} \tag{3-9}$$

$$S_{i,t_0} = 0 \tag{3-10}$$

$$B_{k,t} = B_{k-1,t-\Delta_{k,k-1}} - \sum_{j=1}^{J} d_{j,k} X_{j,t-\Delta_{k,k-1}} + \sum_{i=1}^{I} g_{i,k} M_{i,t-\Delta_{k,k-1}} \tag{3-11}$$

[3] In this queuing diagram, the *nth* vehicle entering the system may not be the nth vehicle leaving the system. Therefore flow data collected at boundaries of the system do not allow us to track the travel times of individual vehicles since the system is not necessary a First-In-First-Out (FIFO) system.

Relationships (3-5) to (3-11) complete the formulation of a linear programming form of the optimal ramp control problem (LP). Inequality (3-6) states that the flows of all freeway sections are not allowed to exceed capacity at any time. Inequality (3-7) is a physical restriction which states that metered flow rates must be positive and not larger than the current on-ramp demands. Equation (3-8) describes the relationship among off-ramp flows, mainline flows and off-ramp diversion rates. Equation (3-9) updates the on-ramp demands and (3-10) is the initialization. Equation (3-11) is a spatially iterative process where metered flow rates *M*, the control variable, finally come into the objective function after substitutions. Although equations (3-11) and (3-8) look like two simultaneous equations, they are actually not if the spatially iterative process starts from the furthest upstream section *K*. Both constraints, inequality (3-6) and (3-7), are non-linear. The non-linearity is introduced by the spatially iterative processes and time dependencies described by equations (3-8), (3-9) and (3-11) when substituted into (3-6) and (3-7).

This new LP does not require any OD information. In fact, the OD information required by other models is substituted by off-ramp diversion rates in this model. Although diversion rates may be determined by an OD trip table, they are also directly measurable by the loop detectors. This measurability enables us to optimize the freeway system without a detailed time-sliced OD table, which is very difficult to estimate.

In formulating the new LP, it is also implicitly assumed that off-ramp diversion rates in any time interval *t* are independent of control decisions (on-ramp flow *M*) made in previous time intervals. This assumption could introduce inaccuracies in the global optimal ramp control formulation described by (3-5) to (3-11) since the causal effects between on-ramp metering rates and off-ramp diversion rates are totally ignored (the exact relationship between on-ramp metering rates and off-ramp diversion rates are determined by the OD demand table and the free-flow travel time matrix *D*). However, in the real time version of this new formulation, the inaccuracies introduced by this assumption will be averaged out since off-ramp diversion rates are updated in each control interval from detector readings. The point is — it is the causal effects between on-ramp metering rates (controlled variables) and off-ramp diversion rates that must be known to optimally control freeways via ramp metering (i.e. minimize total travel time). These causal effects can be obtained from a time-sliced OD trip table, which is the method seen in many previous studies. One can also derive them by knowing the on-ramp metering rates (decisions one makes) and measuring off-ramp diversion rates on a synchronized horizon in real-time, which is the approach taken in this study. There is no need to derive these causal effects from an OD trip table, which is difficult to estimate, when they are directly measurable.

Therefore, we will focus on the real-time version of the new formulation of the global optimal ramp control problem instead of further pursuing solutions to the global version. Also, one has to develop a real-time version of any optimal ramp control theory to implement it since metering-rate decisions in a control interval can only be based on the knowledge accumulated before that control interval (future information can be predicted, however, it is impossible to accurately predict all input variables for the entire control period in a system as dynamic as a freeway).

3.3 A Real-Time Version of the LP

A real-time optimization just optimizes the system in the control interval $t + 1$ (or longer, sometimes called control horizon) based on what one knows at t. A rolling-synchronized-horizon technique is adopted and the global optimization objective function, Equation (3-5), then becomes:

$$Max(B_{t^s} + \sum_{j=1}^{J} X_{j,t^s}) \quad at \quad t^s - 1 \tag{3-12}$$

We use superscript s to denote the synchronized time. If the synchronized time t^s at section 1 is the absolute time t, then the synchronized time t^s at section k is absolute time $t - D_{k,1}$. In other words, at each absolute time t, we determine the ramp metering rates for metered ramp i at absolute time $t - D_{i,1}$. Thus, if one wants to start the control period at absolute time t_0, one must start collecting real-time traffic data at absolute time $t_0 - D_{K,1}$. We call this a rolling-synchronized-horizon objective function. Constraints (3-6) and (3-7) still hold in synchronized time:

$$B_{k,t^s} < C_{k,t^s} \tag{3-13}$$

$$0 \le M_{i,t^s} \le S_{i,t^s} \tag{3-14}$$

In transforming Equation (3-5) to (3-12), dependencies between time slices are thrown out. This ignorance of time dependence allows us to develop a solution process without any predictive elements. If there are any benefits to restricting entering vehicles beyond what would be considered optimal during a single time interval, to allow for more efficient usage of the network in later time intervals, these benefits will be lost in the above transformation. However, it is reasonable to believe that this potential loss of benefits in the time-decomposed version (Equation (3-12)) should not be significantly large due to the following two reasons:

- Constraint (3-6) assures that there are no queues on freeway mainline sections.[4] Therefore, any queues that remain in the system at the end of a control interval and will be dealt with during later control intervals are on-ramp queues.

- The structure of Equation (3-5) tells that allowing more vehicles to leave the system sooner is preferred to restricting them, since the weights on earlier departures are higher than later ones (the weight is (T - t)).

3.4 A Heuristic Solution

We will, in the following paragraphs, develop an intuitive solution to this rolling-synchronized-horizon optimization program. Then we will argue that this solution very likely gives the maximum system departure rate of all possible solutions.

---

[4]Rigorously they could be some transient queues on freeway mainline sections. However, these transient phenomena will have no effects on the control decision at any time interval.

$$(1)\ B_{k,t^s} < C_{k,t^s} \qquad (2)\ 0 \le M_{i,t^s} \le S_{i,t^s}$$

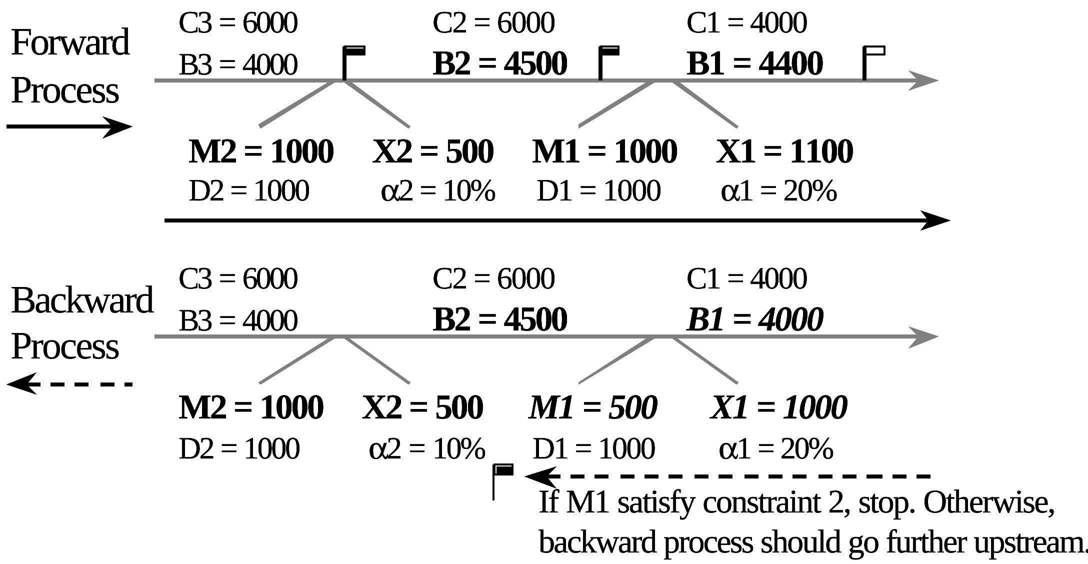


Figure 3-3 A Numerical Example to Illustrate the Heuristic Solution Process

To facilitate the presentation of the heuristic solution, we will use a numerical example in figure 3-3. The heuristic solution consists of repeating a forward process and a backward process (by “forward”, we mean the same direction as traffic flows). Basically, the forward process reflects the nature of the optimization objective function, while the backward process is required by the two constraint equations. In the graph, regular fonts stand for either real-time data collected by detector or threshold capacity values preset by the controller. Bold and italic fonts denote values determined by the forward process and the backward process respectively. Again, all values in the graph are synchronized-time values. We start the “forward process” (the solid arrow) from the furthest upstream section 3 and let all vehicles waiting at the on-ramps (1 and 2) enter the freeway ($M_1 = S_1$, $M_2 = S_2$, the right half of constraint 2, M_i £ S_i, is guaranteed to be satisfied since all the following steps can only decrease the metering rates). Then given these metering rates, we predict the flow of the downstream section (e.g. predict B_2 = 4500 based on B_3, M_2 and a_2). Whenever this forward process proceeds to a new downstream section, we check whether constraint 1 is satisfied (e.g. when the forward process goes to section 2, we check if $B_2 < C_2$. A black-flag in the figure means it is satisfied, a white flag otherwise). When in a section, constraint 1 is not satisfied, it becomes a critical section and a "backward process" will be applied starting from this section and going upstream (the dashed arrow). The backward process first adjusts the on-ramp metering rate at the nearest on-ramp (on-ramp 1) to satisfy constraint 1 at the critical section ($M_1 = 500$). The resulting M_1 at this time could be negative (A negative M_1 means that more than one on-ramp need to be restricted). Then we check whether the left half of constraint 2, $0 \le M_i$, is satisfied at the adjusted on-ramp (on-ramp 1). If it is satisfied, we stop the backward process and start a new forward process from the critical section. Otherwise (negative M_1), the backward process needs to go further upstream (restrict more on-ramps to satisfy both constraints).

The above heuristic solves the rolling-synchronized-horizon version of the LP since all constraints are satisfied. We will then show that this heuristic solution is the global optimal solution to the numerical example. The solution we have now is $M\ (M_1 = 500,\ M_2 = 1000)$. We denote another solution as $M'(M_1', M_2')$. Since M' is a solution to the LP, it must satisfy all the constraints that apply (this means $M_2' \le M_2$ since $M_2 = S_2$). Now there are only three possible types of M':

(1) $M'(M_1' \le M_1$ *and* $M_2' \le M_2$, *but* ~ $(M_1' = M_1$ *and* $M_2' = M_2))$;

(2) $M'(M_1' > M_1$ *and* $M_2' < M_2)$;

(3) $M'(M_1' > M_1$ *and* $M_2' = M_2)$;

It is obvious that (3) does not satisfy constraint 1 in freeway mainline section 1, so it is not a legal solution. Also, under (1) the resulting total departure rate of the system is less than that under M and hence (1) is an inferior solution to M.

Whether M is the global solution now become a comparison between $M'(M_1' > M_1$ *and* $M_2' < M_2)$ with M. If $M_1' = M_1 + n$, we get $M_2' \le M_2 - n/a_2$ to satisfy constraint 1 in section 1, where n is a positive number. Then we can obtain the total system departure rate under M and M':

$Q = B_1 + X_1 + X_2$

Under M: $Q_M = C_1 + C_1(a_1/(1 - a_1)) + B_2 a_2/(1 - a_2)$

Under M': $Q_{M'} \le C_1 + C_1(a_1/(1 - a_1)) + B_2 a_2/(1 - a_2) - n a_2/(1 - a_2)$

Therefore: $Q_{M'} < Q_M$

Thereby $M'(M_1' > M_1 \text{ and } M_2' < M_2)$ is also an inferior solution to *M*. Therefore *M* is the global optimal metering rates to this numerical example. Since our derivation does not depend on the concrete values in the numerical example, this heuristic solution should also be the global optimal solution to this example network with any different parameters. The essential philosophy of the heuristic solution is that, by metering the nearest on-ramp to the critical freeway section, one minimizes the efficiency loss at off-ramps upstream of the critical freeway section. For a larger network, to show *M* is the global maximum, the logical reasoning we just showed in the numerical example is a much more complex task because the number of possible solutions increases exponentially. But we believe that *M* is still the global optimal solution since the philosophy it takes still applies in any larger network.

This heuristic solution has very important qualitative meanings. It states that the most efficient ramp metering control logic is the one metering the nearest upstream entrance ramp(s) to any critical freeway section so as to keep the flow of this section strictly below capacity. However, this most-efficient algorithm is expected to be the least equitable one since the least number of on-ramps will be controlled to provide free-flow conditions for commuters accessing the freeway via other uncontrolled on-ramps.

The heuristic is also very desirable from a computational feasibility perspective. Firstly, it only uses information that has been accumulated to the decision point and no prediction is required. Secondly, the computation work involved in the heuristic is just a straightforward iterative process with simple mathematical operations (only plus and minus) in each iteration step. Finally, all input variables to the solution process are directly measurable from currently available traffic detection hardware. These nice properties of the heuristic make it possible to directly implement it. A ramp control algorithm that implement it will be introduced in Chapter 5. In the next chapter, we will develop an analytical framework for ramp metering studies based on the heuristic.

4 AN ANALYTICAL FRAMEWORK FOR RAMP METERING

The linear program (LP) developed in the previous chapter is a formulation of optimal ramp control problem without internal queues on the freeway (i.e. freeway mainline sections are operated at free flow conditions). Many existing or proposed ramp control strategies are of this kind. There is also an array of ramp metering strategies based on macroscopic traffic flow models and internal queues may be allowed by those strategies[5]. It has been shown that, in the time-independent ramp control problem, preventing the formation of internal queues is a necessary condition that the optimal solution must satisfy (Wattleworth, 1967). However, in the time-dependent cases the benefits of allowing internal queues are not clear. The analytical framework that will be developed herein is only suitable for ramp metering controls not allowing freeway internal queues. Most existing and proposed ramp control strategies belong to this type.

4.1 The Analytical Framework

The heuristic solution to the LP suggests that the most efficient ramp metering control logic is the one **metering the nearest entrance ramp(s)** to **any** critical freeway section so as to **keep the flow of this section strictly below capacity**. In implementing this logic, one has to specify the three[6] following elements and then add these elements to the above logic to form a complete real-world ramp metering algorithm:

4.1.1 Threshold Values – the "capacity" of each freeway section

One must specify the "capacity" of critical sections. In setting these values, one can be either risk averse or risk seeking. Since freeway breakdown is essentially a probabilistic phenomenon (Persaud et al. 1998), one can be risk averse and set critical values in the lower tail of the breakdown probability distribution to minimize the

[5] Whether those ramp control strategies allow internal queues or not can only be tested by comparing the resulting freeway flows with preset freeway capacities and we do not have access to those data.

[6] One can add some optional elements, e.g. smoothness of metering rates. But since these optional constraints are not crucial to the problem, we will focus on these three elements to simplify the discussion. A section will discuss those optional and practical constraints later in this chapter.

probability of failing to satisfy constraint 1 (i.e. freeway mainline section flows must be strictly lower than capacities). However, these somewhat smaller critical values may lead to an efficiency loss. On the other hand, a set of higher critical values can be used with a higher risk of freeway breakdown. This risk-seeking strategy may in the long run improve the overall efficiency of the controlled freeway system.

The "capacity" here can be either flow thresholds or density threshold depending on the control method one chooses (see the next section 4.1.2). In the case of flow capacity, most of the existing ramp control strategies tend to adopt long-run freeway queue discharging flow rates. The threshold set in this way should be considered as risk-averse decisions, not necessarily optimal values. However, it will remain impossible to take advantage of current risk management theories to the case of ramp metering until the probabilistic nature of freeway breakdown is more thoroughly studied. Some studies have provided some insights, though not enough, in this regard (Persaud et al. 1998, Zhang 2002).

4.1.2 Control Methods – how one "keeps the flow of this section strictly below capacity"

The heuristic solution states that the most efficient ramp control strategy must keep the flow of critical sections below critical values. This is a standard control problem and thereby a control method must be selected. In the case of ramp control, one must specify several control details:

Flow control vs. Density control:

One can either control flow or control density to achieve the control objective. Earlier ramp control strategies control flows, mostly because the final control variables, on-ramp metering rates, are actually flows, not density. But recently there is an increasing trend of controlling densities. In density control, some rules must be specified to transform densities to final control variables (not necessarily numerically, e.g. fuzzy rules).

Real-world examples:

Flow control: Twin Cities, Seattle Bottleneck algorithm;

Density control: SWARM, Denver, ALINEA.

Feedback control vs. Feed-forward control:

If there is a discrepancy between the actual flow/density and the desired threshold values, the controller must take action to eliminate/minimize this difference. If the controller employs a predictor to estimate potential discrepancies and take action before the discrepancies actually occur, this type of control is feed-forward. On the other hand, feedback controllers only adjust control parameters (metering rates) based on the detected differences (already occurred) of the desired threshold values and the observed.[7] In general, a feed-back controller is more desirable since feed-forward control does not guarantee convergence. However, to ascertain which type of controller is more suitable for ramp metering, future comparison studies are required.

Real-world examples:

Feedback control: ALINEA, ANN, Denver

Feed-forward control: Twin Cities, most local control algorithms

Linear controller vs. Non-linear controller:

Once a difference term between the desired critical value and the observed/predicted value is identified, will this difference be transformed to control parameters linearly or non-linearly by the controller? For example, in flow control, the simplest controller is the one that directly uses the difference as the metering flow which is a linear controller (*control parameter = difference*).

Real-world examples:

Linear control: ALINEA, Twin Cities, Denver

Non-linear control: ANN (Artificial Neural Network, e.g. Zhang 1997)

[7] A controller using a Kalman filter (e.g. SWARM algorithm) is sort of a mix of both feed-forward and feedback controllers since the Kalman filter contains both a prediction equation to predict future system states and a feedback equation to reduce prediction errors based on measured data.

4.1.3 Equity Considerations

Coordinating on-ramp meters is often a necessary step to satisfy constraint Equation (3-6). However, it can also be viewed as an equity consideration. A theoretical way to consider equity in the optimal ramp control problem would be to reset the optimization objective from minimizing total absolute travel time to minimizing total weighted travel time (ramp delays have higher weights than free flow travel times).

The optimization objective function in the formulation of the optimal ramp control problem in Chapter 3 is to minimize total absolute travel time. This objective given constraint Equation (3-6) is essentially the same as minimizing total ramp delays, because once the freeway is operated at free-flowing condition and the demand is fixed, total freeway mainline travel time will become a constant term which can be taken out of the objective function:

total absolute travel time = total freeway mainline travel time + ramp delays

if (total freeway mainline travel time = Constant):

$$\textit{Max(total absolute travel time)} = \textit{Max (ramp delays)} \qquad (4\text{-}1)$$

Equation (4-1) actually tells us that if ramp delays are weighted linearly (all ramp delays have the same value of time), then our solution developed in Chapter 3 is also the solution for minimizing the total weighted travel time. This means that a linear value of time for ramp delays is not only unrealistic but also useless in balancing efficiency and equity. Therefore, one has to apply appropriate non-linear weights to ramp delays (i.e. longer delays have heavier weights).

Some practical equity considerations have also evolved over time in real-world ramp control strategies. The first control constraint that improves ramp control equity is probably the maximum queue length restriction, although the original purpose of this constraint is to prevent ramp queue spillover to local streets. In all practical ramp control strategies, there is a minimum/maximum metering rate constraint which is also beneficial from an equity point of view. More specifically, the Denver strategy has a so-called "helper algorithm" among on-ramps in which if one ramp is operated at its most restrictive rates, the ramp that immediately upstream of it will be operated more restrictively in the next control interval to release the downstream one to some extent. The Minnesota Zonal algorithm controls all on-ramps in a control zone (called co-ordination groups in some other algorithms) to assure the flow at the zone bottleneck is below capacity. The new Minnesota Stratified Zonal algorithm has a maximum ramp delay restriction which ensures that the maximum ramp delay will not exceed four minutes.

4.1.4 Why This Analytical Framework Is Useful

Up to now, we have completed the construction of an analytical framework under which many existing/proposed ramp metering algorithms can be viewed as ramifications of one ramp control logic. The differences among those algorithms are that they have different threshold values and/or different control methods and/or different equity considerations. Under this framework, these elements that constitute ramp control strategies can be easily decomposed and studied individually. Previous studies have tested different ramp metering algorithms in simulators. One algorithm can be shown better than the other under the same simulation scenario by comparing simulation outputs. However, most such studies are unable to answer why the more efficient algorithm is more efficient. Hence, these studies are generally nothing more than engineering comparisons with very limited theoretical implications. This happens because there are so many elements in each algorithm and each of these elements can make a difference of the performance of the whole algorithm. The analytical framework developed above provides us a way to decompose these individual factors and hence comparisons can focus on only one of various factors keeping all others equal, which can not only give us more valuable results but also a more generic research area. A research topic under this framework could be, for instance, "All others being equal, can a non-linear controller provide a more efficient results than a linear controller" or, "Provided the same threshold values and control method, how will the efficiency of the control strategy trade off with equity considerations", or "Which equity consideration is most effective in improving the equity of the controlled freeway system".

4.2 Some Practical Issues Regarding Ramp Metering

Although many mathematical theories and techniques have been used in solving the optimal ramp control problem over years, the nature of ramp metering has never changed – it is an engineering problem. Like many other engineering problems, theories cannot cover every aspect of the optimal ramp control problem. Sometimes, these practical aspects not addressed in the theory are actually very important to the success of a project and hence deserve a discussion in this report.

Data Smoothing Factor

It is a common experience that traffic data collected by inductive loop detectors or some other detection facilities need to be smoothed to deal with both transient traffic fluctuations and minor detection errors, especially when the control interval is short (e.g. 30 seconds). The impact of smoothing factors and smoothing methods (moving averages, exponential smoothing, etc.) on the performance of ramp metering strategies (or other traffic control systems) has not been systematically studied. However, there is no reason to believe this impact can be simply neglected.

In the previous section, we show that the proposed analytical framework can decompose various elements that are contained in a complete ramp metering algorithm, but the potential impacts of smoothing factors are not mentioned. A profile optimization method could be used to account for this problem. For example, if one wants to compare the effectiveness of flow control and density control, one could use various smoothing factors for both control scenarios respectively and the comparison should be made between the best flow control scenario with its optimal smoothing factor and the best density control scenario with its optimal smoothing factor.

Detector Malfunctioning

For people who have been working with data collected by loop detectors, it is not a surprise to see malfunctioning detectors. How to improve the reliability of loop detectors is obviously out of the scope of this report. But a good ramp control strategy should not be very sensitive to the readings of one single detector. This is a control reliability issue. The ability of an algorithm to deal with malfunctioning detectors should be tested before its implementation unless the reliability of loop detectors themselves are significantly improved.

Non-recurrent Freeway Congestion

It has been reported that more than 50% of freeway delays result from non-recurrent freeway congestions such as traffic accidents. Ramp metering control algorithms are not explicitly designed to relieve non-recurrent congestion, nevertheless an algorithm that can also deal with non-recurrent congestion is more desirable.

5 CO-EOAX RAMP METERING ALGORITHMS

In this chapter, the heuristic solution introduced in Chapter 3 will be implemented. As mentioned in Chapter 4, the ramp metering logic suggested by the heuristic must combine several more elements to form a complete ramp control strategy. Efficiency Oriented Algorithm (EOA) is a direct implementation of the heurist with risk-averse threshold values, flow control, feed-forward control, linear controller and no equity consideration.

Since the major goal of this study is to develop a both efficient and equitable ramp control strategy, EOA obviously dose not meet the goal. Therefore, Co-EOAX (Coordinated Efficiency Oriented Algorithms with global grouping factor X[8]) will be developed. When X is equal to 1, Co-EOA1 is simply EOA. As X increases, the equity of the controlled freeway system will also increase at the expense of efficiency. The methodology to determine the optimal X value will be presented in the next Chapter.

5.1 Efficiency Oriented Algorithm (EOA)

Efficiency Oriented Algorithm (EOA) is a direct implementation of the heuristic with risk-averse threshold values, flow control, feed-forward control, linear controller and no equity consideration. It has been coded in C++ and a flowchart of it is shown in Figure 5-1. One may note that the flowchart is quite different from the forward/backward processes we introduced in Chapter 3. The fact is that the repeated forward/backward processes actually can be implemented very easily under the help of a so-called "exchanger" transient variable (*Ex* in the flowchart; it is called "exchanger" since this variable just exchanges information between adjacent freeway sections)[9].

---

[8] On-ramp global grouping factor X is the number of immediately upstream on-ramps that are to be coordinated to release a critical freeway section, i.e. a freeway section where the total flow is about to exceed the capacity. However, X = 1 does not mean there is no coordination at all and the actual grouping factor for each final coordination group is not always equal to X, which will be explained in details in appendix B of this report.

[9] Since the EOA is coded based on GETRAM Extension and CPI (see section 1.2.3 for related background information) and these two function libraries are under the protection of their copyrights, the EOA source code can not be published along with this report. Interested readers may contact the author for related information.

In each control interval, the EOA control algorithm is implemented as follows (the notation is the same as those in section 3.1):

Step 0: Start from the furthest upstream freeway mainline section $k = 1$ and the exchanger (Ex) is set to a large value (larger than the capacity of any freeway section).

Step 1: Identify the type of section k from the freeway coding input file (to be introduced later). A freeway section can only be one of three types (metered section, unmetered section or offramp section).

Get current control threshold value CR_k (the minimum of the capacity of section k and the current exchanger Ex):

$$CR_k = min\ (C_k, Ex)$$

Step 2:

if section k is an metered section:

Update upstream section upstream flow C_k;

Update maximum possible ramp flow S_i (on-ramp i is on section k);

Update HOV bypass flow U_i (if there is a HOV bypass).

Compute the control variable – metered flow M_i:

$M_i = CR_k – B_k - U_i$

if $Mi < 0$:

$Ex = CR_k + M_i$

$M_i = 0$

else if $M_i > S_i$:

$M_i = S_i$

$Ex = large$

else if section k is an unmetered section:

$Ex = CR_k - U_i$

else section k is an offramp section (offramp j is on section k):

$$Ex = CR_k/(1 - a_j)$$

Step 3: Pass M_i (if any) to the meter controller.

Step 4:

if section k is the last section on the controlled freeway:

end;

else:

$k = k + 1$

Ex = current Ex

Go back to step 1.

The input file required by the EOA is a freeway-coding file in the following format (the sample is the file for Trunk Highway (TH) 169 northbound from I-494 to Medicine Lake Road in the Twin Cities, MN, USA):

section id: 1
section type: 1
smoothing factor: 2
default section capacity: 37
upstream station name: 768
metered ramp name: 3C7
departure station name: 3061
arrival station name: MedicineLake
HOV station name: none
……
section id: 15
section type: 2
smoothing factor: 2
default section capacity: 100
upstream station name: 439
unmetered ramp name: HOV4
departure station name: 1932
……
section id: 35
section type: 3
smoothing factor: 2
default section capacity: 100
upstream station name: 426

offramp station name: 1352

Where:

Section id: the index of a freeway mainline section (see section 3.1 for the basic freeway coding convention). The furthest downstream section of the controlled freeway is section 1.

Section type: the type of a section. 1 stands for a metered section, 2 for unmetered section and 3 for offramp section.

Smoothing factor: the data smoothing factor for a section (see section 4.2).

Default section capacity: C in our notation, the unit is vehicle per 30 seconds.

Upstream station name: the name of the detection station that provide upstream flow rates of a section.

Unmetered station name: if a section is an unmetered section, the name of the unmetered ramp meter (if any).

Departure station name: the name of the station that provides departure flow rates of the on- ramp (metered or unmetered) on this section.

Offramp station name: if a section is an offramp section, the name of the detection station that provides exiting ramp flow rates of the offramp on this section.

Metered ramp name: if a section is a metered section, the name of the meter.

Arrival station name: if a section is a metered section, the name of the detection station that provides arrival flow rates of the on-ramp on this section.

HOV station name: if a section is a metered section and there is a HOV bypass on this section, the name of the detection station that provides HOV departure rate of the on-ramp on this section ("*none*" means there is no HOV bypass).

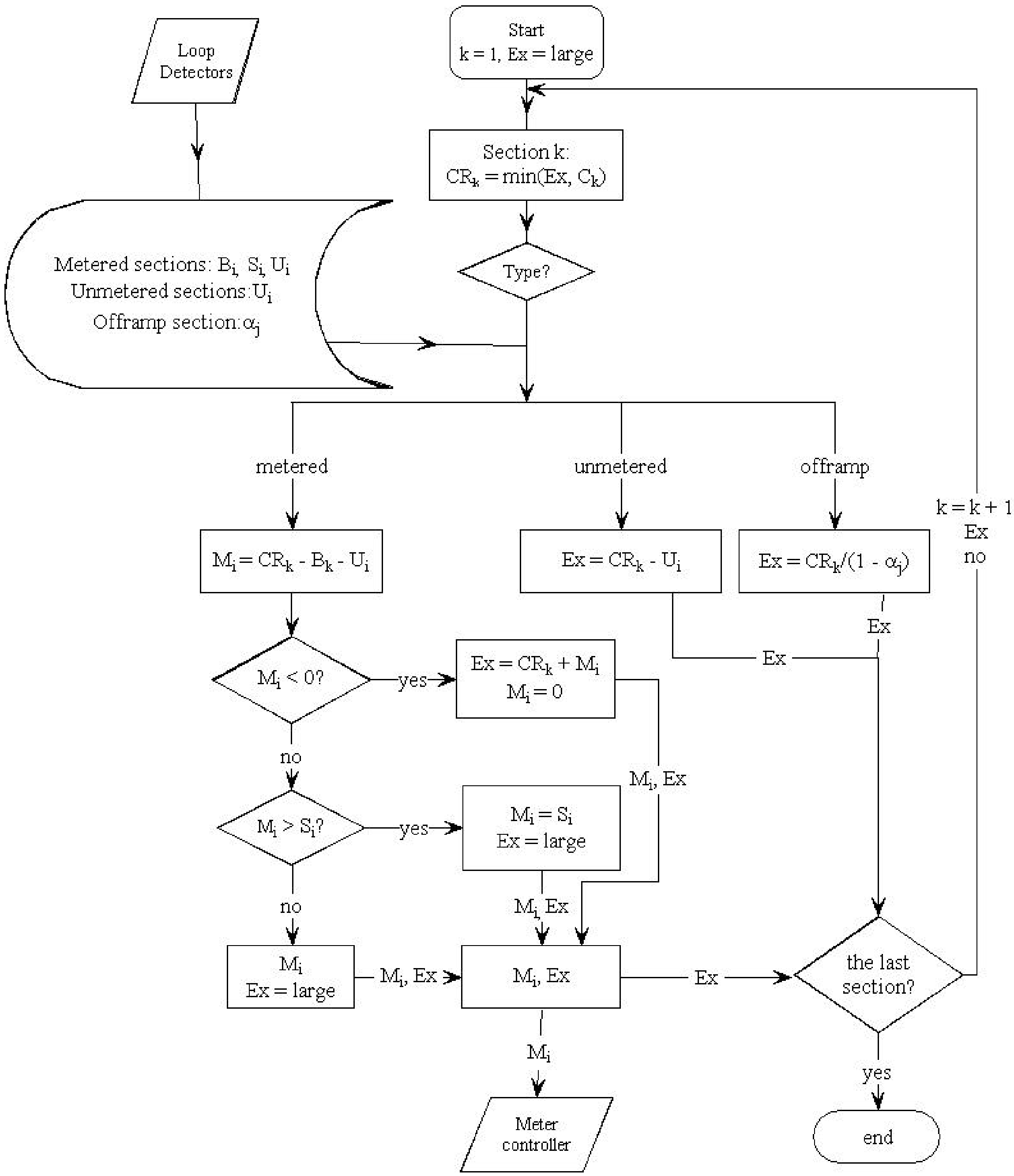


Figure 5-1 Flowchart of EOA

5.2 Coordinated Efficiency Oriented Algorithm with Global On-Ramp Grouping Factor X (Co-EOAX)

The EOA is simply the implementation of the heuristic developed in Chapter 3. Before we proceed to the details of the Co-EOAX, it is necessary to introduce the motivation for developing Co-EOAX first. In section 4.1.3, it is shown that one has to apply non-linear weights to ramp delays to effectively incorporate equity considerations into the optimal ramp control problem. However, the linear program (LP) formulated in section 3.2 is based on the freeway system level queuing analysis and hence individual ramp delay is not identifiable. To formulate an optimal ramp control problem which aims to minimize total non-linearly weighted travel time, a set of on-ramp First-In-First-Out (FIFO) queuing systems must be introduced to capture individual ramp delays. This will tremendously increase the complexity of the already-complex linear program, because a set of transformation equations converting on-ramp departure/arrival rates to individual ramp delays must be added and a set of non-linear weights will be introduced into the optimization objective function. An alternative, much simpler way is to first define a vector of equity-control variables which explicitly consider equity. Here, a vector of global on-ramp grouping factors, ***X****s*, will be used as equity variables (see footnote 1 of this chapter). Each control variable X gives a ramp metering control strategy which is named Co-EOAX (**Co**ordinated **– EOA** with global on-ramp grouping factor **X;** how the coordination among on-ramps is achieved is detailed in the next section). Then all these control strategies ($X = 1, 2, 3, \ldots$) can be simulated and the total weighted travel times of all strategies can be computed using simulation outputs. The optimal equity control variable X_{opt} corresponding to the control strategy Co-EOAX$_{\text{opt}}$, which minimizes the total weighted travel time, thus can be determined.

The mechanism of Co-EOAX is illustrated using the example of Co-EOA3, the global on-ramp grouping factor X = 3. Again, note that all the values in the following presentation are synchronized values. The mechanism of Co-EOAX is, in short, a combination of on-ramp coordination processes and freeway mainline section checking processes (see figure 5-2):

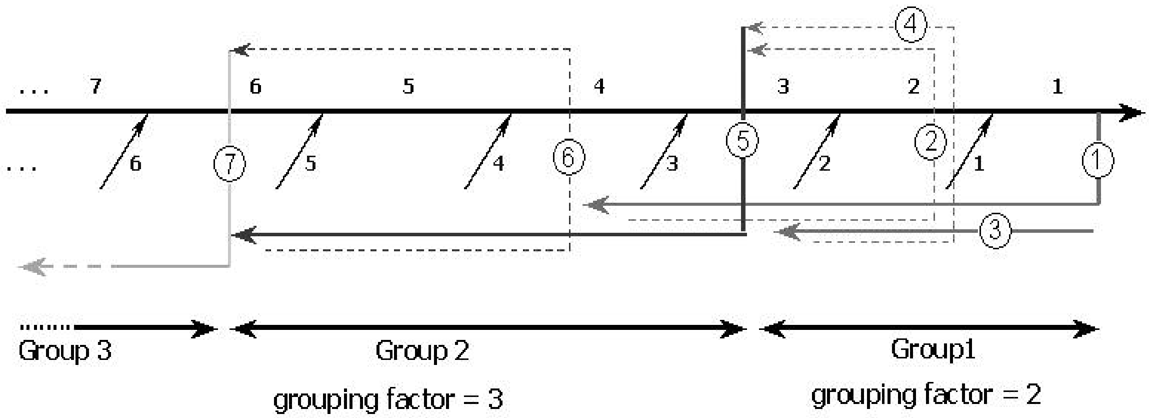


Figure 5-2 Co-EOA3

Step 0: update detector measurements at all freeway sections;

Step 1 (On-ramp coordination): starting from the furthest downstream section (section 1). Section 1 is considered a critical section (it does not matter whether it really is one or not). The critical value for section 1 is just its capacity C_1. The three ramps upstream of it will be coordinated to ensure the flow in section 1 will not exceed its capacity since the global on-ramp grouping factor $X = 3$. All these three ramps are coordinated such that the ramp delays at these ramps are the same. This is achieved by equalizing the ratio of the metering rates to the maximum possible ramp flows at these on-ramps, $M_1/S_1 = M_2/S_2 = M_3/S_3 = R$. If $R > 1$, R is set to be 1 (section 1 is actually not a critical location) and the "exchanger" Ex is set to be a arbitrary large value. If $R < 0$, R is set to be 0 and the Ex is set to be $C_1 + R(S_1 + S_2 + S_3)$. If $0 < R < 1$, $Ex = large$. The ramp metering rates for the three ramps are S_1R, S_2R and S_3R respectively (these metering rates are temporary rates and may not be equal to the final rates, see the following steps). If Ex is not equal to "large", this Ex will affect the coordination in the next coordination group, see step 5. In essence, Ex is a parameter which enables ramp coordination beyond the current coordination group. However, this coordination enforced by Ex is in order to satisfy constraint 1 at critical freeway sections, not an equity consideration, which makes it different from the coordination enforced by the global grouping factor X.

Step 2 (Checking): Given the metering rates determined in step1, predict the freeway section flow between the current critical section (section 1) and the furthest upstream on-ramp (on-ramp 3), i.e. the current coordinated freeway sections (section 2 and section 3). Then using the predicted section flow to check whether constraint 1 is satisfied at all the current coordinated sections starting from the downstream section (section 2). Section 1 does not need to be checked since constraint 1 at this section is automatically satisfied in step 1. The first section where constraint 1 is not satisfied will become the critical section for the next coordination group. In the figure, constraint 1 is not satisfied in section 3 and hence section 3 will become a new critical section. Also, the coordination we did in step 1 for the old critical section (section 1) needs to be redone with a new grouping factor.

Step 3 (On-ramp coordination): Regrouping on-ramp 1 and 2 since on-ramp 3 is no longer in this coordination group and the new grouping factor is 2. Compute the new R, the new metering rates ($M1, M_2$) for on-ramp 1 and 2 and the new Ex.

Step 4(Checking): Check all coordinated sections in the coordination group with critical section 1 and grouping factor 2. In this case, check constraint 1 in section 2 only. Since constraint 1 is satisfied in section 2, we can now proceed to coordinating the group with the new critical section – section 3. The metering rates determined in step 3 will then be the final rates for on-ramp 1 and 2.

Step 5 (On-ramp coordination): The coordination process in step 5 is the same as that in step 1 except for two things:

1. We are now coordinating on-ramps upstream of section 3.

2. Since section 3 is not the furthest downstream section, the critical value for section 3 will be taken as the minimum of the capacity of section 3 and the current Ex value. Since the global grouping factor is 3, we will coordinate three on-ramps upstream of section 3. We compute the R, metering rates (M_3, M_4 and M_5) and the Ex.

Step 6 (Checking): We now check constraint 1 at all current coordinated sections (section 4 and 5). Since constraint 1 at section 4 and 5 is satisfied. We now move to the next critical section – section 6.

Step 7 (On-ramp Coordination): The critical value for section 6 is the minimum of the capacity of section 6 and the current *ex*. We then coordinate three on-ramps upstream of section 6 since global grouping factor is 3.

The on-ramp coordination and checking processes in Co-EOA3 are fully described in the above seven steps. These coordination and checking processes will be repeated until the metering rates for all on-ramps in the controlled freeway system are determined. If the global grouping factor is not equal to 3, we just use this new global grouping factor in the above steps. It should be pointed out again that the final grouping factor of each on-ramp coordination group is not always equal to the global grouping factor (e.g. for group 1, the final grouping factor is 2, not 3). This is because the ramp coordination enforced by equity considerations sometimes conflicts with the constraint that requires the flow of each freeway section below its capacity.

The input file required by the Co-EOAX is also only a freeway-coding file in the following format (the sample is still the file prepared for the TH169 northbound from I-494 to Medicine Lake Road in the Twin Cities):

section id: 1
section type: 1
smoothing factor: 2
default section capacity: 37
upstream station name: 768
metered ramp name: 3C7
departure station name: 3061
arrival station name: MedicineLake
HOV station name: none
coordination section names:1,3,5,7,9,11,13,17,19,21,22,24,26,28,30,32,34
… …
section id: 15
section type: 2
smoothing factor: 2
default section capacity: 100

upstream station name: 439

unmetered ramp name: HOV4
departure station name: 1932
……
section id: 35
section type: 3
smoothing factor: 2
default section capacity: 100
upstream station name: 426
offramp station name: 1352

The only difference between the EOA input file and the Co-EOAX input file is that there is one more line for metered sections in the Co-EOAX input file listing the ids of all possible on-ramp coordination sections delimited by commas. For instance, if X = 3, for section 1, section 1, 3 and 5 would be the metered sections that will be coordinated in the first coordination process.

6 MINIMIZING WEIGHTED TRAVEL TIME

Both the field experiment results evaluated in Chapter 2 and the optimal ramp control theory developed in Chapter 3 suggests that efficiency trade off with spatial equity. The purpose of this chapter is to develop a methodology to identify an optimal ramp control strategy which minimizes weighted travel time of the controlled freeway system, so as to balance efficiency and equity in a systematical way. As stated in section 4.1.3, the weights applied to travel times (delays) must be non-linear weights, otherwise the new objective of minimizing weighted travel time would be essentially the same as minimizing absolute travel time when freeway internal queues are not allowed. In section 5.2, the difficulties of directly minimizing non-linearly weighted travel time are briefly introduced and the idea of an alternative simulation method is discussed.

This simulation method will be detailed in this chapter. Co-EOAXs are coded in C++ based on GETRAM Extension and CPI (see section 1.2.3). The first section illustrates the scheme of how the Co-EOAX algorithms and AIMSUN2 microscopic traffic simulator interact. The next section presents the format of the simulation output provided by the AIMSUN2 simulator and how the total system weighted travel time can be extracted from the simulation outputs. The final section describes a real-world freeway system on which Co-EOAX ramp control strategies will be tested in the AIMSUN2 simulator.

6.1 Simulating Co-EOAXs in the AIMSUN2 Simulator

The information exchanges between the Co-EOAX algorithms and the AIMSUN2 simulator are realized by two intermediate interfaces: GETRAM Extension and CPI. The process of the information exchanges is illustrated in Figure 6-1. In each simulation step, the AIMSUN2 model of the road network emulates the traffic detection process. Then, through a set of functions in the GETRAM Extension, it provides external applications with "simulation detection data" (flow, occupancy etc.). CPI is a module that contains some functions that are especially useful for external ramp control strategies. The simulation detector data is processed and aggregated to the desired level in CPI. The Co-

EOAX algorithms read data from CPI and determine the metering rates for all control ramp meters on the simulated freeway. These control actions are fed back to CPI and then GETRAM Extension. The AIMSUN2 model can only read external control decisions from the GETRAM Extension and then emulate the ramp control operations though the corresponding model components.

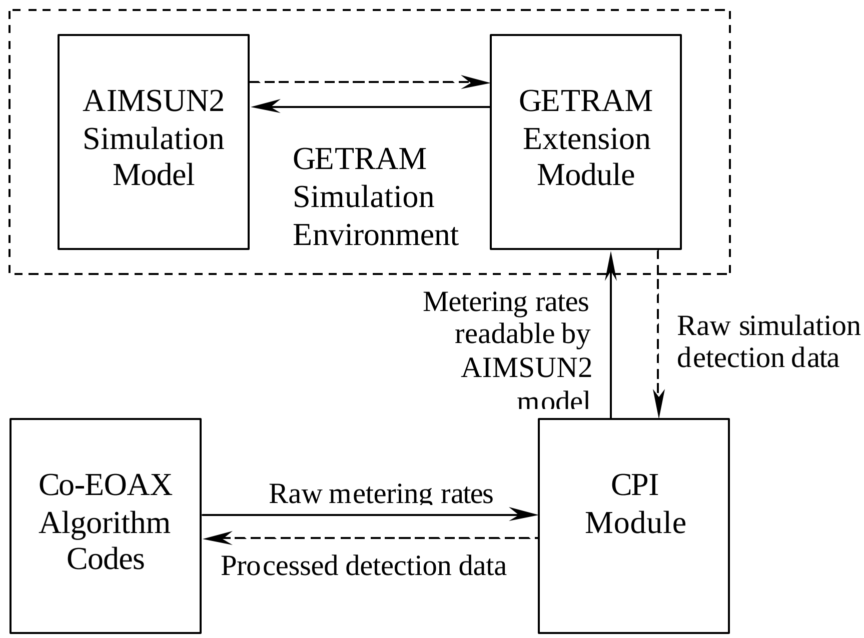


Figure 6-1 Scheme of Simulating Co-EOAX Algorithms

6.2 Extraction of Total Weighted Travel Time from AIMSUN2 Simulation Output

The AIMSUN2 microscopic simulator provides three types of output: animation of the simulation, detection data and statistical results. The last two types of output will be used to obtain total weighted travel times under Co-EOAX ramp control strategies.

Detection data can be collected in the simulator at user-specified intervals. 30-second detection intervals will be used in this study, which is consistent with the detector data aggregation intervals in the Twin Cities.

The statistical traffic measures provided by AIMSUN2 can be specified at different levels of aggregation: for the whole system, for each section, for each turning movement, for every stream (set of consecutive sections) defined by the user. The statistical measures can be presented according to two time scopes: (1) Global – statistical data gathered from the beginning to the end of a simulation experiment; (2) Periodic – statistical data gathered during user-specified time periods (5 minutes in this study). The statistical measurements provided by the AIMSUN2 simulator in each collection period include flow rates, density, mean speed, mean travel time, mean delay time, number of stops etc. However, the AIMSUN2 simulator does not provide system weighted travel time directly. Therefore, a methodology to extract total weighted travel time from AIMSUN2 simulation output must be developed.

The total absolute system travel time *ATT* can be obtained from the statistical output:

$$ATT_X = \frac{\sum_{r=1}^{R}\sum_{t=1}^{T}\sum_{k=1}^{K}\left[\frac{t_{k,t}Q_{k,t}}{12}\right]}{3600R} \qquad (6\text{-}1)$$

Where:

ATT: absolute total system travel time (vehicle*hours);

X: index of Co-EOAX algorithms (e.g. $X = 1$ corresponds to Co-EOA1);

r: index of simulation replications;

R: total number of simulation replications (same for all control strategies);

t: index of 5-minute statistical data collection periods;

T: total number of data collection periods;

k: index of sections (freeway mainline sections and ramp sections) in the network;

K: total number of sections;

$t_{k,t}$: average travel time per vehicle in section *k* during the t^{th} 5-minute period (statistical measurements provided by the AIMSUN2 output) (seconds);

$Q_{k,t}$: average flow through section *k* during the t^{th} 5-minute period (statistical measurements provided by the AIMSUN2 output) (vehicle per hour).

Then individual ramp delays will be computed using a queuing analysis at on-ramps. This queuing analysis is the same as the one introduced in section 2.2. The input data for the queuing analysis are on-ramp arrival and departure rates collected by the AIMSUN2 simulator. The computed ramp delays, written in a vector form, are:

$$d_X(d_{1,1,X}, d_{2,1,X}, \ldots, d_{v,1,X}, \ldots, d_{V_{1,X},1,X}; \ldots\ldots; d_{1,r,X}, d_{2,r,X}, \ldots, d_{v,r,X}, \ldots, d_{V_{r,X},r,X};$$
$$\ldots\ldots; d_{1,R,X}, d_{2,R,X}, \ldots, d_{v,R,X}, \ldots, d_{V_{R,X},R,X}) \tag{6-2}$$

Where:

$V_{r,X}$: total number of vehicles entering the freeway network via metered on-ramp throughout the simulation period in the r^{th} replication of the simulation experiment under control strategy Co-EOAX;

$d_{v,r}$: ramp delay experienced by the v^{th} vehicle in the r^{th} replication of the simulation experiment under the corresponding control strategy Co-EOAX, (seconds);

Total absolute ramp delay (*ARD*, veh*hrs) is just the sum of individual ramp delays:

$$ARD_X = \frac{\sum_{r=1}^{R} \sum_{v=1}^{v=V_{r,X}} d_{v,r,X}}{3600R} \tag{6-3}$$

Total weighted ramp delay (*WRD*, veh*hrs) is the sum of non-linearly weighted individual ramp delays:

$$WRD_X = \frac{\sum_{r=1}^{R} \sum_{v=1}^{v=V_{r,X}} f(d_{v,r,X})}{3600R} \tag{6-4}$$

Where:

f(.): a non-linear weighting function with slope larger than 1 anywhere in the range (0, +8).

The slope of the non-linear weighting function must be larger than 1 in the range (0, +8), which ensures that ramp delays have higher weights than free flow travel times (weights = 1).

Although non-linear travel time has been explored by many previous studies, unfortunately none of them provides a non-linear weighting function for ramp delays (similar to the concept of "stop/start travel time" in these studies). Some studies answer questions of this type – "Is *n* minutes delay as onerous as *n* minutes free flow travel time? If not, what is the elasticity between delayed travel time and free flow travel time". For instance, Hensher (2000) estimates the elasticity between stop/start time and free flow time from multinomial logit and mixed logit models using data collected from a stated preference survey. The mean stop/start time is 2.6 minutes and the resulting elasticity ranges from 8 to 17. These studies do estimate non-linear weighting functions for travel times. However, these weighting functions are non-linear only globally and if one narrows the range to delayed time only (or stop/start time only) that portion of these functions is still linear (see Figure 6-2). It is shown in section 4.1.3 that these globally

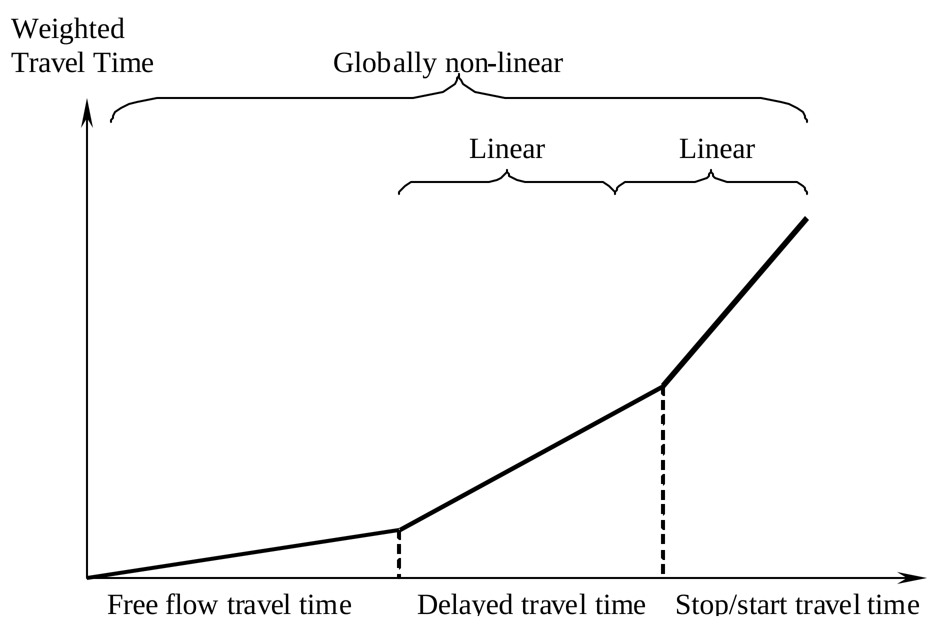


Figure 6-2 Weighting Functions Only Non-linear Globally

non-linear value-of-time functions can not help balance the efficiency and equity of a ramp control strategy. We here simply assume a plausible form for the ramp delay weighting function *f(.)* in which the ramp delay weight increase as ramp delay itself increases but the weights increase at a diminishing rate (see Figure 6-3):

$$f(d_{v,r}) = \begin{cases} 4d_{v,r} & \text{if } d_{v,r} < 30 \text{ seconds;} \\ 8d_{v,r} & \text{if } d_{v,r} < 120 \text{ seconds;} \\ 16d_{v,r} & \text{if } d_{v,r} < 300 \text{ seconds;} \\ 20d_{v,r} & \text{Otherwise.} \end{cases} \qquad (6\text{-}5)$$

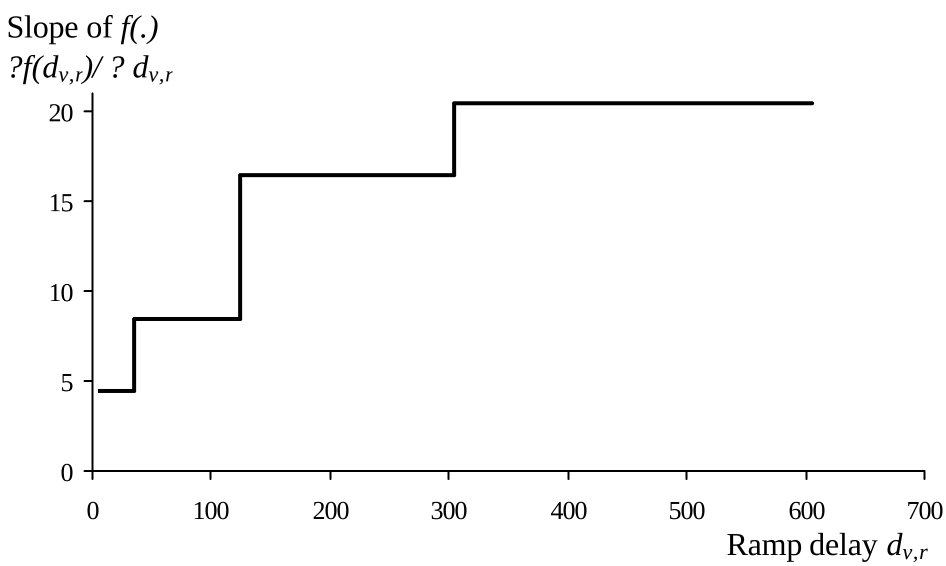


Figure 6-3 A Plausible Non-Linear Ramp Delay Weighting Function

Substituting absolute ramp delay (*ARD*) from total absolute travel time (*ATT*), we obtain freeway mainline travel time (*FTT*, veh*hrs):

$$FTT_X = ATT_X - ARD_X \qquad (6\text{-}6)$$

Since all Co-EOAX algorithms do not allow freeway mainline queues, the freeway mainline travel times under these algorithms are all free flow travel times. Hence, the weighted freeway mainline travel time (*WFTT*, veh*hrs) is equal to absolute freeway mainline travel time (*FTT*) since the weight for free flow travel time is one:

$$WFTT_X = FTT_X \qquad (6\text{-}7)$$

Now, the total weighted travel time (*WTT*, veh*hrs) is simply the sum of the weighted ramp delay (*WRD*) and the weighted freeway mainline travel time (*WFTT*):

$$WTT_X = WRD_X + WFTT_X \qquad (6\text{-}8)$$

The optimal global on-ramp grouping factor, X_{opt}, then can be determined and Co-EOAX$_{opt}$ is the algorithm that minimizes total weighted travel time of all Co-EOAX algorithms:

$$WTT_{X,MIN} \Rightarrow X_{opt} \qquad (6\text{-}9)$$

Where:

$WTT_{X,MIN}$: the minimum weighted system travel time of all Co-COAX algorithms.

Equation (6-1) through (6-9) completes the methodology of extracting total weighted system travel times from simulation output provided by the AIMSUN2 simulator. Next, all Co-EOAXs will be tested on a real-world freeway system in the simulator, not only to demonstrate the above methodology of identifying Co-EOA$_{opt}$, but also to evaluate the efficiency and equity of the resulting Co-EOA$_{opt}$.

6.3 A Test Freeway Section

One of the four freeway sections evaluated in Chapter 2 – TH169 northbound from I-494 to I-94 will be used to test Co-EOAXs (Figure 6-4). This freeway site is chosen because it has already been calibrated in the AIMSUN2 simulator in an earlier study (Hourdakis and Michalopoulos 2002). Most of the test section consists of two lanes with ten weaving sections. It has 24 entrance ramps of which one is unmetered. The metered ramps include 4 HOV bypasses and two freeway-to-freeway ramps from TH62 and I-394. The test site contains 25 exit ramps. The upstream and downstream boundaries are free of congestion when the traffic demand data of this section were collected on March 21, 1999.

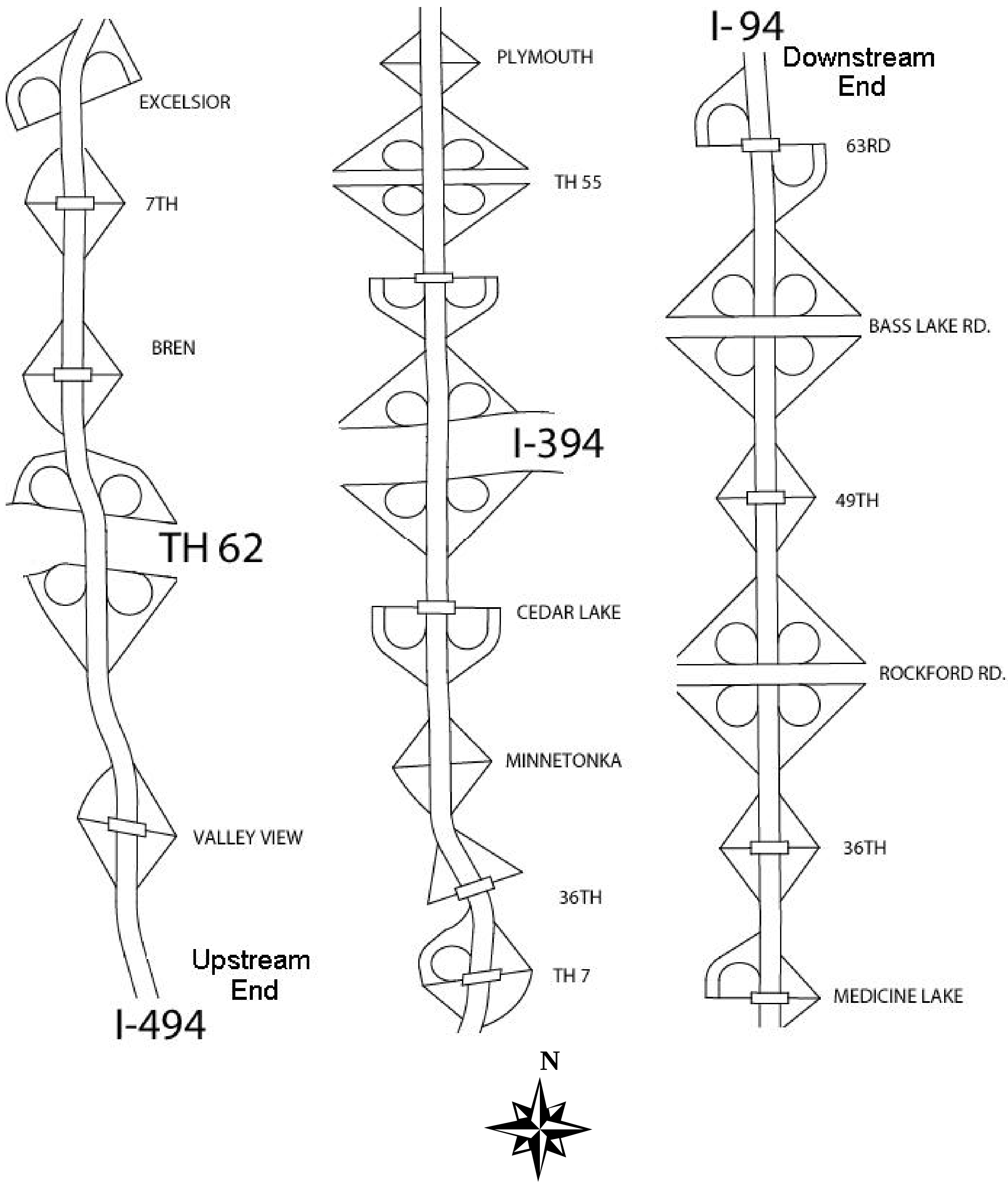


Figure 6-4 TH169 Northbound Test Freeway Section

7 SIMULATION RESULTS ON THE TEST SITE

The coded Co-EOAXs are simulated on the TH169 northbound network. Total system absolute travel time, total absolute ramp delay, freeway mainline travel time and total system weighted travel time for each control strategy are computed using the method developed in Chapter 6. In this simulation scenario (i.e. the freeway network and the traffic demand data), the optimal global on-ramp grouping factor is 2 ($X_{opt} = 2$). Hence, Co-EOA2 is the optimal control strategy that minimizes total weighted travel time. Metering-off case and the Minnesota algorithm are also simulated under the same simulation scenario to compare with the Co-EOAXs.

The efficiency and equity of Co-EOA2, Minnesota algorithm and no-control case are further evaluated using the MOEs developed in Chapter 2. Co-EOA2 is found to be both more efficient and more equitable than the Minnesota algorithm.

7.1 Identifying the Algorithm Minimizing Total Weighted Travel Time – Co-EOAX_{opt} on the Test Site

Five simulation runs have been replicated for each ramp control scenario (including the metering-off case) with the same five random seeds. The results are summarized in Table 7-1 and illustrated graphically in Figure 7-1 through Figure 7-4. The results support the optimal ramp control theory developed in Chapter 3, 4 and 6:

- Co-EOA1 (or EOA) is the most efficient ramp metering algorithm which minimizes total absolute travel time. The total absolute travel time is 10 percent less under the Co-EOA1 control compared to the no-control scenario. The Minnesota algorithm only shortens the total travel time by 7 percent.
- As X increases (or equity is considered more and more), the system efficiency drops. When X is larger than 2, Co-EOAXs become less efficient than the Minnesota algorithm. More over, when X is larger than 5, the resulting system efficiency is even worse than the no-metering case.

- When the metering objective is to minimize total weighted travel time, Co-EOA1 is no longer the optimal ramp control strategy. In this test scenario, Co-EOA2 is the most desirable control strategy that minimizes total weighted travel time.
- The total weighted travel time under the no-control scenario can not be computed since freeway mainline queues will form without ramp metering. In that case, a weighting function for freeway mainline delays needs to be developed.
- Freeway mainline travel times among Co-EOAXs are almost the same since the freeway mainline is operated under free flowing conditions and the traffic demand is fixed for all Co-EOAXs.
- Co-EOA1 results in shortest total ramp delay. As X increases, the total ramp delay also increases, but the total weighted ramp delay could decrease. In this case, the total weighted ramp delay under Co-EOA2 is smaller than that under Co-EOA1.

Table 7-1 Summary of Simulation Result on TH169 Northbound Test Site

Unit: veh*hrs	AbsoluteTotal Travel Time	Weighted Total Travel Time	Freeway Mainline Travel Time	Total Ramp Delay
Co_EOA1	6240	16523	5839	401
Co_EOA2	6321	15702	5766	554
Co_EOA3	6471	18848	5758	713
Co_EOA4	6657	21714	5799	857
Co_EOA5	6795	24302	5789	1006
Co_EOA6	6948	27918	5749	1199
Minnesota	6412	15949	5756	661
No Control	6929	N/A	6929	0

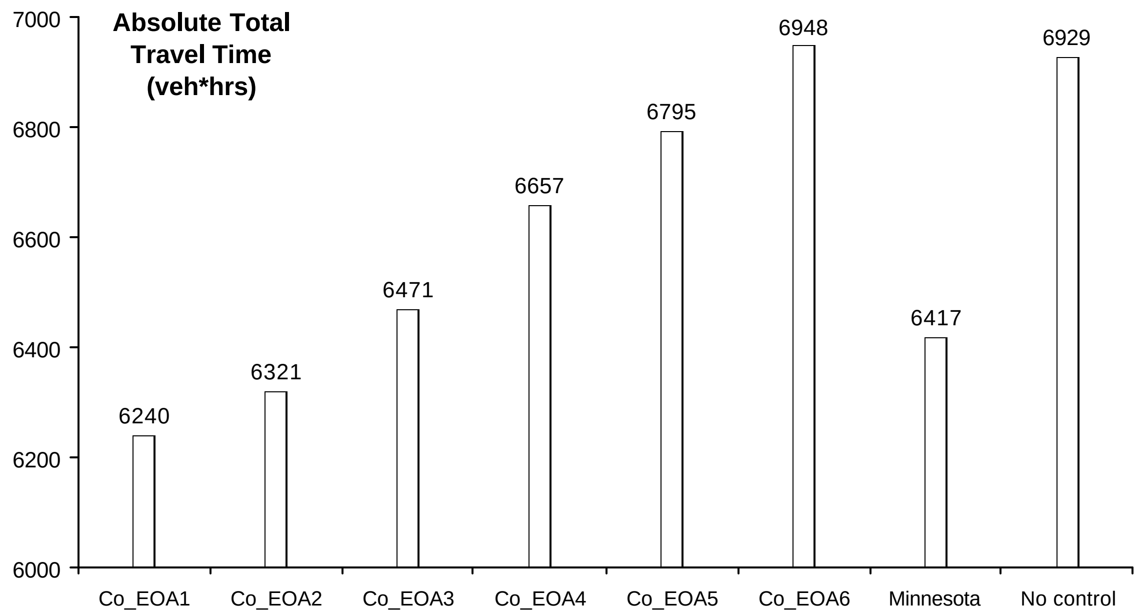


Figure 7-1 Absolute Total Travel Time

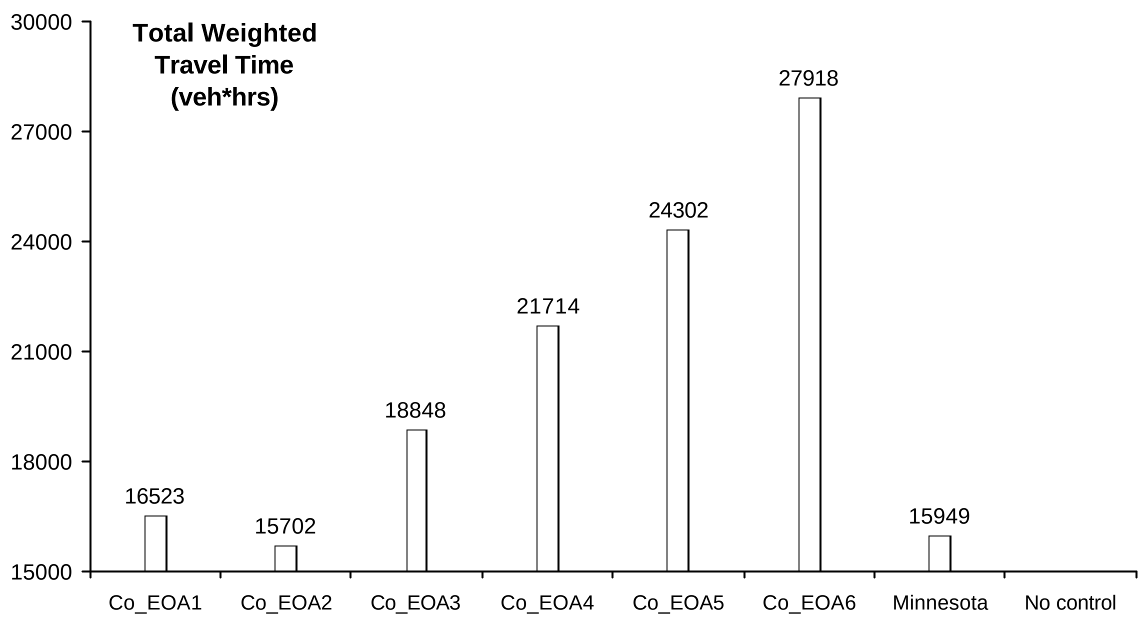


Figure 7-2 Weighted Total Travel Time

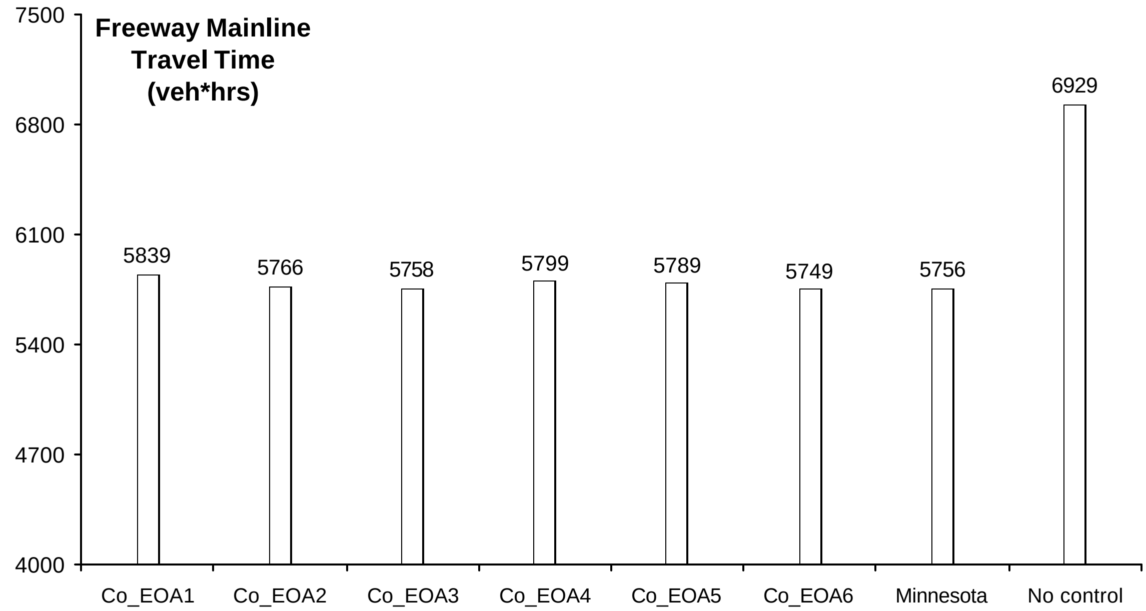


Figure 7-3 Freeway Mainline Travel Time

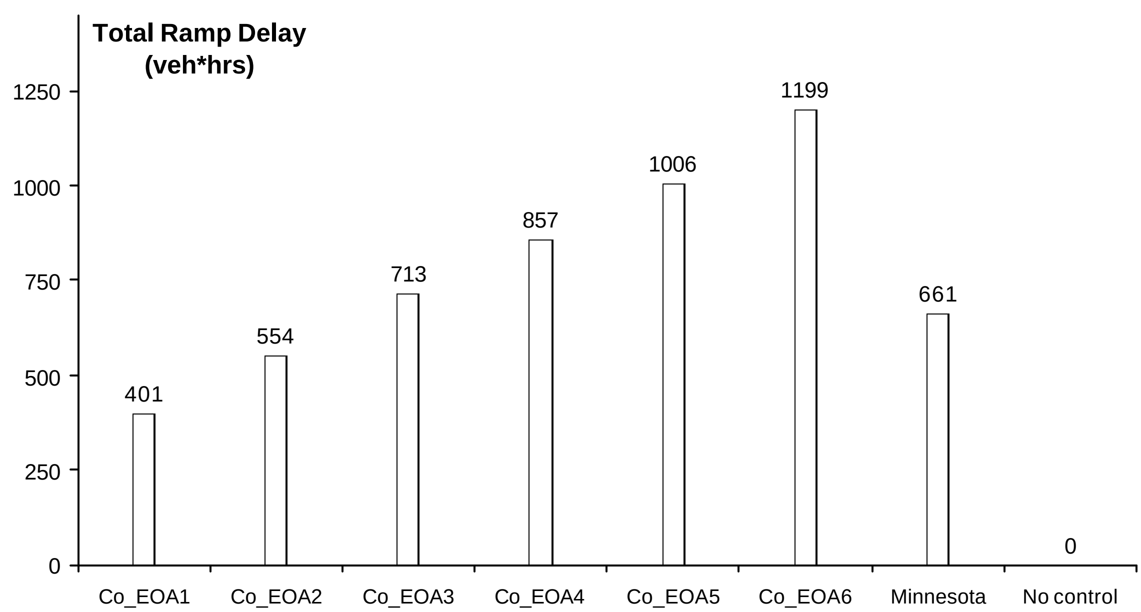


Figure 7-4 Total Ramp Delay

7.2 Evaluating Efficiency and Equity of Co-EOAX$_{opt}$ and the Minnesota Algorithm

The optimal global on-ramp grouping factor suggested by the simulation results in the previous section is 2. To determine whether Co-EOA2 provides better balance between efficiency and equity, or is both more efficient and more equitable than the Minnesota algorithm, these two ramp control strategies are evaluated using the efficiency and equity measures developed in Chapter 2. The evaluation methodology is the same except now the simulated detector data will be used instead of field detector data. Results are shown in Figure 7-5.

When the peak is approached, Co-EOA2 tremendously improves the overall mobility of the freeway system compared to the no-control case, while the system spatial equity is still comparable to that of the no-control case. Under the Minnesota algorithm, although the length and severity of congestion are reduced compared with no-control, this reduction is at a considerable expense of equity. Clearly, Co-EOA2 is more efficient and more equitable than the Minnesota algorithm.

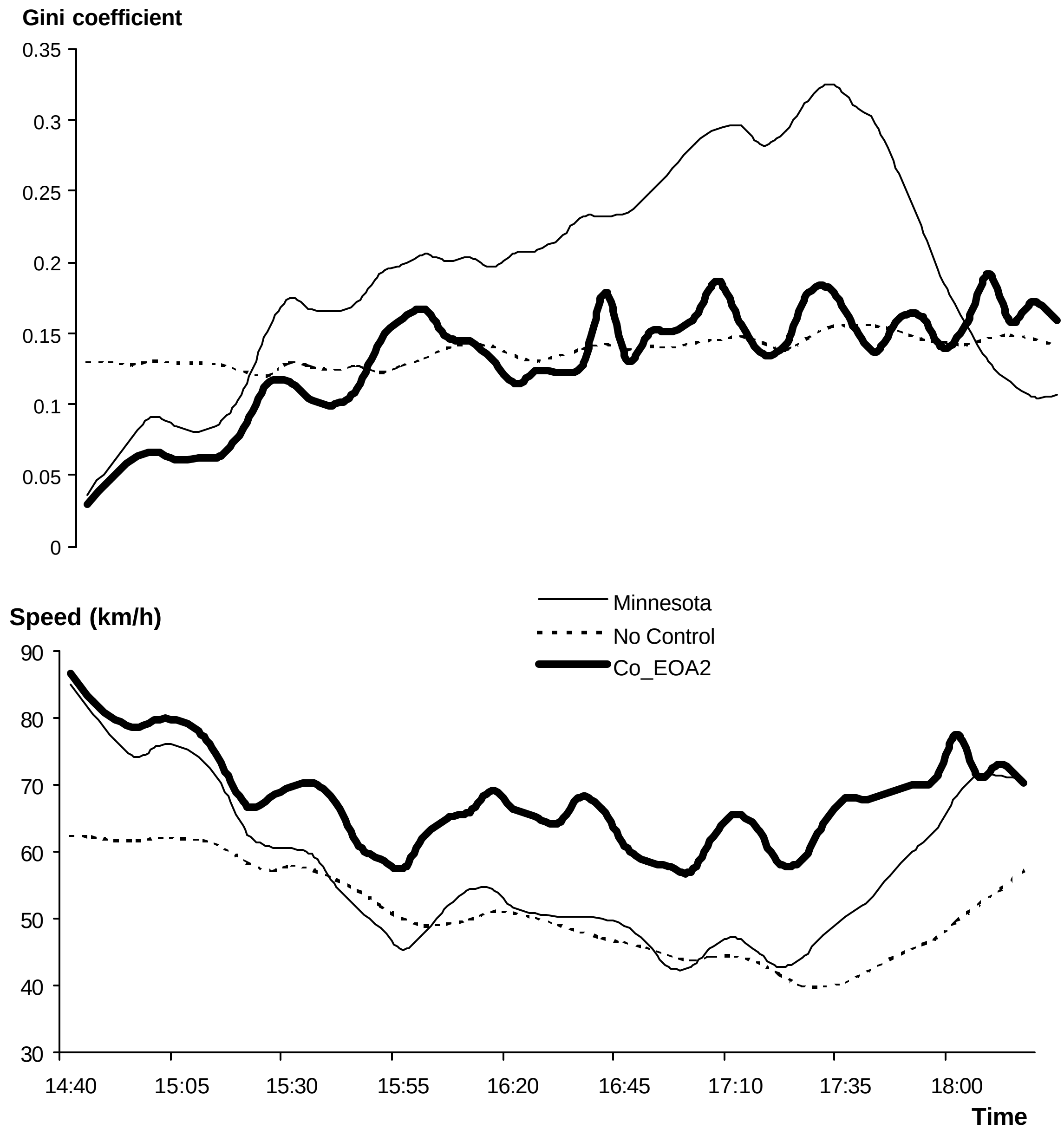


Figure 7-5 Efficiency and Equity of Co-EOA2, the Minnesota Algorithm and No-Control

8 CONCLUDING REMARKS

This study developed a new theory of efficient and equitable ramp metering control. The study has developed measures of efficiency and equity of ramp meters, formulated the optimal ramp control problem without requiring origin-destination (OD) information estimation, provided a heuristic solution to it, created an analytical framework under which future ramp metering research can be conducted, developed and coded a new family of ramp metering algorithms with different equity consideration, and introduced a new ramp metering objective and shown how it can be used to balance efficiency and equity of ramp control strategies.

A methodology to measure efficiency and equity of a freeway system, as well as some other important performance indicators (accessibility, productive, consumers' surplus, travel time variation and demand responses), is developed in Chapter 2. These performance measures can be routinely computed on any freeway systems with traffic detectors installed since the only input data required by the methodology are volumes and occupancies collected by traffic detectors, which is useful in evaluating impacts of any kinds of control policy changes or demand shifts (e.g. traffic growth) on freeway system performance. As an example, this methodology is used in this study to evaluate effectiveness of a real life ramp metering control strategies. The evaluation methodology has been coded in a Python script and can be readily used by a Freeway Traffic Management Center or other parties.

Some findings in evaluating the effectiveness of the real life ramp control strategy – the Minnesota algorithm, deserve to be highlighted: (1) a trade-off between efficiency and spatial equity is evident; (2) improvements in freeway travel reliability could be a huge benefit of ramp metering and thus should be captured in future benefit/cost studies; (3) Some demand responses to ramp metering hypothesized by previous studies are confirmed.

Previous formulations of the ramp optimal control problem require OD demand information and are all encountered by the difficulty to estimate accurate real-time OD information. Also, the presence of OD information in these previous formulations complicates the problem and makes it almost impossible to obtain global optimal solutions. Although most theoretical studies on ramp metering assume the availability of time-sliced OD information, interestingly, it has never been proved that knowing this information is a necessary condition to optimally control freeways via ramp metering (i.e. minimize total system travel time). In Chapter 3, a new formulation of continuous-time optimal ramp control problem without OD information is described. All input variables for the optimization program in the new formulation are directly measurable by loop detectors. The real-time version of the new control formulation is also developed and a heuristic solution to it is presented. The heuristic is demonstrated to be the global optimal solution on a small-scale network. Future work should pursue a mathematical proof of a global optimal solution to the new optimization program.

The heuristic solution implies that the most efficient ramp control strategy is also the least equitable one, which provides a qualitative foundation for the development of an analytical framework for ramp metering. Under this analytical framework, all ramp control strategies not allowing freeway mainline queues can be viewed as ramifications of the same metering logic suggested by the heuristic, with different critical values, control methods and equity considerations. Then it becomes possible to decompose individual elements of a complete ramp metering control strategy and then to study them separately. Future studies can be conducted under this framework focusing on only one of various factors affecting the overall effectiveness of ramp control strategies keeping all others equal. In the long run, such studies should provide more valuable results than those directly comparing two existing or proposed ramp control strategies.

Since all information required by the heuristic solution is directly measurable by loop detectors in real-time, it is possible to implement it directly on freeways where detectors have been installed. The Efficiency Oriented Algorithm (EOA) that implements the heuristic is developed. It is the theoretically most efficient ramp control strategy without any equity consideration. Co-EOAX (Coordinated Efficiency Oriented Algorithm with global on-ramp grouping factor X) has also been developed which are derived from EOA but with more equity considerations. This new series of ramp control algorithms not only are developed on a solid theoretical foundation, but also can be implemented in real world without any additional data collectors.

The Twin Cities ramp metering shut-off experiment illustrates the emerging need to balance efficiency and equity of ramp control policies. Ramp control strategies developed by traffic engineers usually aim to improve freeway efficiency or mobility (e.g. minimize total travel time). Improved mobility for the system as a whole is a good thing. However, if that is achieved by helping some drivers at the expense of others, there is also a serious equity issue that should be considered. Co-EOAXs are desirable under this background since the balance point between efficiency and equity can be changed to the direction decision-makers prefer simply by increasing or decreasing the value of a single parameter – global on-ramp grouping factor X.

Simply minimizing total travel time on a freeway system should not be the ultimate goal of ramp metering since this objective results in strategies that are purely efficiency-oriented. As an alternative, a new metering objective – minimizing total weighted travel time is introduced. When delays on the freeway system are weighted non-linearly (i.e. longer delays are weighted more), equity considerations are actually built into this objective. In this study, a simulation method to minimize total weighted travel time is developed. This methodology is demonstrated using Co-EOAXs on a sample real-world freeway network in AIMSUN2 microscopic traffic simulator. In this scenario, Co-EOA2 is the optimal ramp control strategy, judged by the new objective. Using measures of efficiency and equity developed in Chapter 2, Co-EOA2 is also found to be both more efficient and more equitable than the Minnesota algorithm.

A theoretical way to minimize total weighted travel time requires reformulation of the optimal ramp control problem using this new objective. Also, a need for reasonable non-linear value-of-travel-time function is also identified. These two subjects are left to future research.

REFERENCES

Barceló, J., Ferrer, J.L., and Grau, R. (1994). *AIMSUN2 and the GETRAM simulation environment*. Internal report. Departamento de Estadística e Investigación Operativa. Facultad de Informática. Universitat Politècnica de Catalunya.

Bates, J., Dix M., and May, A.D. (1987). Travel Time Variability and its Effect on Time of Day Choice for the Journey to Work. *Transportation Planning Methods, Proceedings of Seminar C*, PTRC. pp. 299-311.

Black, I.G., and Towriss J.G. (1993). Demand Effects of Travel Time Reliability. Center for Logistics and Transportation, Cranfield Institute of Technology.

Bogenberger, K. and May, A.D. (1999). *Advanced Coordinated Traffic Responsive Ramp Metering Strategies.* California PATH Working Paper, *UCB-ITS-PWP-99-19.*

Cambridge Systematics (2001). *Mn/DOT Ramp Metering Study Final Report*. http://www.dot.state.mn.us/rampmeterstudy/ Accessed on November 2004.

Chang, T.H., and Li, Z.Y. (2002). Optimization of mainline traffic via an adaptive co-ordinated ramp-metering control model with dynamic OD estimation. *Transportation Research part C*, Vol. 10, 99-120.

Fonda, R. D. (1969). An analysis of short-term implementation of ramp control on the Dan Ryan Expressway. Abridgement, *Highway Research Record 279.*

Gini, C. (1936). On the Measure of Concentration with Espacial Reference to Income and Wealth. Cowles Commission .

Glasser, G. J. (1962). Variance Formulas for the Mean Difference and Coefficient of Concentration. *J. Amer. Stat. Assoc*., Vol 57, pp. 648-654.

Gordon, D.R., and Wynne, T. (1973). *Evaluation of fixed-time ramp control on the San Diego Freeway in Los Angeles*. *Report 73-4,* California Department of Transportation, Freeway Operation Branch, District 7.

Haj-Salem, H., and Papageorgiou, M. (1995). Ramp Metering Impact on Urban Corridor Traffic: Field Results. Transportation Research Part A, Vol. 29, Issue 4, 303-319.

Hensher, D.A. (2000). The Valuation of Travel Time Savings for Urban Car Drivers:

Evaluating Alternative Model Specifications. California Path ITS-PWP-00-9.

Hourdakis, J., and Michalopoulos, P.G. (2002). Evaluation of Ramp Control Effectiveness in Two TwinCities Freeways. *Tran. Res. Rec.* (In Press).

Koka, M. (2001). Assessing the effectiveness of ramp control on Twin Cities freeways, Master's Thesis, University of Minnesota.

Kotsialos, A., and Markos P. (2001). Efficiency versus fairness in network-wide ramp metering, *2001 IEEE Intelligent Transportation Systems Conference Proceedings*, Oakland (CA) USA, August 25-29, 2001.

Kotsialos, A., Papageorgiou, M., Mangeas, M. and Haj-salem, H. (2002). Coordinated and integrated control of motorway networks via non-linear optimal control. *Transportation Research Part C*, Vol. 10, 65-84.

Kwon E., — (2000). *Comparative analysis of operational algorithms for coordinated ramp metering*, Working Report prepared for the Minnesota Department of Transportation.

Kwon, E., Nanduri, S.V., Lau, R., and Aswegan, J. (2001). Comparative Analysis of Operational Algorithms for Coordinated Ramp Metering. Presented at the 80th Transportation Research Board Conference in Washington DC, January 2001.

Lawson, T.W., Lovell, D.J., and Daganzo, C.F. (1997). Using the Input-Output Diagram to Determine the Spatial and Temporal Extents of a Queue Upstream of a Bottleneck. *Transportation Research Record* 1572.

Levinson, David M. and Kumar, Ajay, (1994). Operational Evidence of Changing Travel Patterns: A Case Study. *ITE Journal* April 1994.

Levinson, D., and Kumar, A. (1995). A Multimodal Trip Distribution Model: Structure and Application. *Transportation Research Record* 1466, pp.124-131.

Levinson, D., and Kumar, A. (1997). Density and the Journey to Work. *Growth and Change*, 28:2, 147-172.

Levinson, D. (2002a). Perspectives on Efficiency in Transportation. Presented at the 81st Transportation Research Board Conference in Washington DC, Jan. 2002.

Levinson, D. (2002b). Identifying Winners and Losers in Transportation. *Transportation Research Record* (in press).

Lomax, T.J. and Schrank, D.L. (2000). *A comparison of ramp metering in Minneapolis/St. Paul and peer U.S. cities*. Working Report prepared for the Minnesota Department of Transportation.

Lorenz, M.O. (1905). Methods of Measuring the Concentration of wealth. *Publication of the American Statistical Association 9*, 1995. pp. 209-219.

Lovell, D.J. (1997). *Traffic control on metered networks without route choice,* Ph.D. Dissertation, University of California – Berkeley.

Lovell, D.J. and Windover, J.R. (1999). Analyzing Freeway Traffic under Congestion: Traffic Dynamics Approach - Discussion by D.J. Lovell and J.R. Windover. *Journal of Transportation Engineering*, July/August 1999.

Lovell, D. and Daganzo, C. (2000). Access control on networks with unique origin-destination paths. *Transportation Research Part B*, 34, 185-202.

May, A. D. (1964). Experimentation with manual and automatic ramp control. *Highway research record* 59.

Minnesota Department of Transportation (1996). *Traffic Management Program Overview: Twin Cities Metro Area.* Report Number *TMC07043-0196* February 1996.

Minnesota Department of Transportation Metro Division, Traffic Management Center, (1998). *Ramp Metering by Zone – The Minnesota Algorithm*. Working Report

Minnesota Department of Transportation (2002a). Accessed on August 2002 http://www.dot.state.mn.us/rampmeterstudy/faqs1010.html#19

Minnesota Department of Transportation (2002b). Accessed on August 2002 http://www.ajfroggie.com/roads/rampmeter.htm

Newell, G.F. (1982). *Application of Queuing Theory*. 2nd ed. Chapman and Hall.

Newman, L., Dunnet, A., and Meis, G. J. (1969). Freeway ramp control - what it can and can not do, *Traffic Engineering.* Vol. 39, pp. 14-21.

Papageorgiou, M., (1983). *Application of Automatic Control Concepts to Traffic Flow Modeling and Control*. Springer-Verlat.

Papageorgiou, M., Hadj-Salem, H., Blosseville, J. (1991). ALINEA: A local feedback control law for on-ramp metering, *Transportation Research Record* 1320.

Papageorgiou M., (1995). An integrated control approach for traffic corridors. *Transportation Research Part C*, Vol. 3, Issue1, 19-30.

Persaud, B., Yagar, S. and Brwonlee, R. (1998). Exploration of the Breakdown Phenomenon in Freeway Traffic. *Transportation Research Record*, paper no. 98-1567.

Pinnell, C., Drew D., McCasland, W., and Wattleworth, J. (1967). Evaluation of entrance ramp control on a six-mile freeway section, *Highway Research Record* 157.

Piotrowicz, G., and Robinsons J. (1995). *Ramp metering status in north America, DOT-T-95-17*, Federal Highway Administration.

Pisarski, A. (1992). *New Perspectives in Commuting*. Washington, DC: US Department of Transportation.

Ries, G. (1980). *Impact of Weather on Freeway Capacity. Mn/DOT Metro Division,* Traffic Management Center research project report (unpublished).

Reiss, R.A., Nathan, H.G., and Stephen, L.C. (1991). Dynamic control and traffic performance in a freeway corridor: A simulation study, *Transportation Research Part A*, Vol. 25, Issue 5.

Small, K.A., Noland, R.B., and Koskenoja, P. (1995). Socio-economic Attributes and Impacts of Travel Reliability: A Stated Preference Approach Draft PATH Report, *MOU-117*.

Small, K.A., Noland, R.B., Chu, X., and Lewis, D. (1999). *Variation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation*. NCHRP Report 431.

Star Tribune Minneapolis (1999). State senator Dick Day offers ‘Freedom to Drive’ agenda, October 18, 1999.

Stephanedes, Y.J., Kwon, E., and Michalopoulos, P. (1992). Control emulation method for evaluating and improving traffic-responsive ramp metering strategies, *Transportation Research Record* 1360.

Taylor, C., and Meldrum, D. (2000). *Evaluation of a Fuzzy Logic Ramp Metering Algorithm: A Comparative Study Among Three Ramp Metering Algorithms Used in the Greater Seattle Area*. WSDOT Technical Report WA-RD 481.2.

Taylor, C., Meldrum, D., Jacobson, L. (1998). Fuzzy Ramp Metering – Design Overview and Simulation Results, *Transportation Research Record* 1634.

Wascoe, D. Jr. (2000). Do ramp meters encourage urban sprawl? *Minneapolis Star Tribune*, Oct. 22, 2000.

Wattleworth, J.A. (1963). *Peak-period control of a freeway system – some theoretical considerations*. Ph.D. dissertation, Northwestern University.

Wattleworth, J.A. (1967). Peak-period analysis and control of a freeway system. *Highway Research Record*.

Weibull, J. W. (1980). On the Numerical Measurement of Accessibility. *Environment and Planning A*, Vol. 12.

Zhang H.M., Richie, S.G. and Recker, W.W. (1996). Some general results on the optimal ramp control problem. *Transportation Research Part C*, Vol. 4, Issue 2, 51-69.

Zhang, H.M. (1997). Freeway Ramp Metering Using Artificial Neural Network. *Transportation Research Part C*, Vol. 5, Issue 5, 273-286.

Zhang, H.M., and Recker, W.W. (1999). On optimal freeway ramp control policies for congested traffic corridors. *Transportation Research Part B*, Vol. 33, 417- 436.

Zhang, L. and Levinson, D. (2003). Does ramp metering exacerbate sprawl? Presented at the 13th ITSA Annual Meeting, Minneapolis, MN, May 2003.

Zhang, L. (2004). Prediction of freeway breakdown using Artificial Neural Network. Presented at the 14th ITSA Annual Meeting, San Antonio, TX, April 2004.