CENTER FOR
TRANSPORTATION
STUDIES

UNIVERSITY OF MINNESOTA

DEMAND-RESPONSIVE RAMP METERING CONTROL TO IMPROVE TRAFFIC MANAGEMENT IN FREEWAY CORRIDORS, PHASE I

Yorgos Stephanedes,
Eil Kwon, Kaikuo Chang,
and Ping Yao,
Civil and Mineral Engineering

Sponsored by:
Center for Transportation Studies

DEVELOPMENT AND APPLICATION OF DEMAND-RESPONSIVE RAMP METERING CONTROL TO IMPROVE TRAFFIC MANAGEMENT IN FREEWAY CORRIDORS

PHASE I

FINAL REPORT

Principal Investigators: Yorgos J. Stephanedes, P.E. and Eil Kwon

Graduate Students: Kaikuo Chang and Ping Yao

Department of Civil and Mineral Engineering

University of Minnesota

January 1992

EXECUTIVE SUMMARY

A method is developed for evaluating traffic-responsive ramp metering strategies and improving freeway performance. The method emulates real-time metering and rigorously traces the interactions between automatic rate-selection metering strategies and freeway performance through time. Given a demand pattern and freeway geometrics, it provides assessment of metering strategies that change continuously at very short time intervals. Further, it explicitly treats time delays that can be caused by hardware or introduced by the traffic engineer.

The method was tested with real data in emulating volume and occupancy thresholds, rate tables and automatic rate-selection control strategies on I-35W in Minneapolis. The tests indicated errors in the range from 5 to 12 percent. The method was also tested in assessing and selecting ramp control strategies based on freeway performance measures such as total freeway volume and delay. Evaluation results were in the expected direction, and indicated the ability to distinguish small performance differences across strategies.

The method can be a practical tool for facilitating traffic management by aiding in the design of new traffic-responsive strategies at individual ramps or sets of ramps, and in the evaluation of existing control systems and selection of the best metering strategies. Using this tool for considering the trade-offs between performance indicators such as volume increase and delay reduction, a desirable threshold policy can be determined for a given freeway section prior to implementation.

Although the new method improves on the performance of conventional systems, it is still restricted by modeling and hardware limitations. For achieving an on-line performance needed in determining optimal metering strategy configurations in real-time large freeway networks, further research is ongoing by the authors. This research addresses the need for improved traffic prediction, parallelization of large-scale computations, and use of improved data collection techniques for the derivation of traffic data. Work is also ongoing for automating the selection of optimal strategies on arterials with the anticipation that the two lines of work will lead to improved selection of demand responsive control strategies in freeway corridors and urban networks.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF TABLES

I. INTRODUCTION

I.1 PROBLEM STATEMENT

Freeway congestion has become a common phenomenon in urban areas leading to delays, reduced traffic safety, increased fuel consumption and severe air pollution. Managing traffic congestion is of critical importance to the preservation of the viability of the urban environment, where traffic jams already cost $73 billion a year to our economy. In the Twin Cities Metro Area, freeway delays are expected to almost triple by the year 2005 (Regional Transit Board, 1986).

With no drastic solutions in sight, the most cost-effective alternative that could immediately address this problem is the optimal use of existing capacity by distributing traffic demand in space and time through efficient on-line management of traffic in a freeway network. The most advanced concept for such an optimal management includes a hierarchical traffic control structure, where overall freeway control is decomposed into several components, such as demand prediction, network optimization and direct control (see Figure 1), so as to achieve computational feasibility and robustness of control solutions. While this concept is promising, the state of the art in freeway ramp metering has not reached the point where a comprehensive, network-wide control solution is automatically generated and implemented through on-line optimization. A major difficulty lies in the lack of accurate on-line predictors, that can predict traffic demand and diversion in freeway networks, and the lack of efficient computational algorithms implementable for on-line optimization.

As a result of the above limitations, most traffic-responsive metering systems, such as the Twin Cities freeway control system, employ automatic rate-selection procedures. These procedures select the most appropriate metering rates for a ramp from a pre-determined library using the information received from loop detectors on the main freeway, upstream (feedforward control) and downstream (feedback control) from the ramp. Although this method provides a degree of self-adjustment to prevailing traffic conditions, the inherent limitations in terms of possible rate configurations and the lack of an efficient procedure for updating threshold values significantly restrict the effectiveness of control. This limitation is especially evident as the size of the metered freeway network increases.

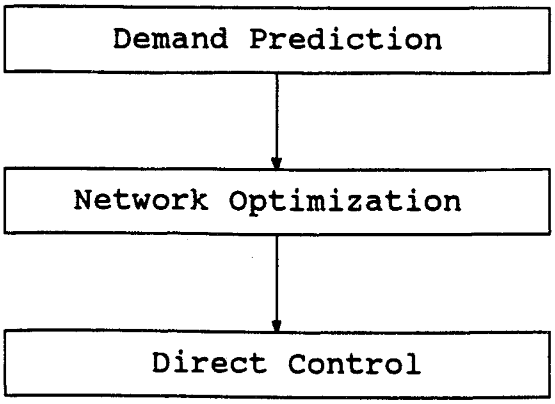

Figure 1. Hierarchical Freeway Control Structure

Development of realistic and practical on-line metering strategies is an essential element in improving traffic management in freeway networks. The control strategies should be based on the predicted behavior of traffic flow responding to the control to be applied in a network, where conditions on any given segment affect many other segments in the network. Further, an efficient parallel processing algorithm, which determines the optimal metering rates for sets of network ramps simultaneously, should be developed to reduce the computational time, so that the control solutions can be implemented in real time.

I.2 PROJECT OBJECTIVES

The ultimate goal of this research is the development and application of on-line optimal ramp metering control strategies in freeway corridors; the strategies will reflect the network-wide traffic conditions and flow interactions through on-line demand-diversion prediction and efficient on-line optimization. The research will be accomplished in 4 phases. The major objective of the current project, which is part of Phase I, is the development of an efficient analytical tool, that can evaluate existing automatic rate-selection strategies, the most common type of metering in operation. The 4 phases are summarized below.

Phase I: Development of a control-emulation method for automatic rate-selection ramp metering systems,

Phase II: Field application of the control-emulation method and preliminary study for the development of on-line optimal control strategies,

Phase III: Development of on-line optimal ramp metering strategies for a small freeway network, and

Phase IV: Development of on-line optimal ramp metering strategies for a freeway corridor and a large urban network.

The major accomplishments of the current project, Phase I, are,

- Review of current freeway metering systems and existing optimal control strategies.
- Development of a control-emulation method for evaluating automatic rate-selection metering strategies.
- Example application of the control-emulation method.

I.3 REPORT ORGANIZATION

The second chapter describes the traffic-responsive metering algorithms currently operating in the U.S. and other countries. Chapter 3 reviews advanced concepts for hierarchical, optimal freeway control presented in the literature. Chapter 4 develops a control-emulation method to evaluate automatic rate selection strategies and Chapter 5 contains test results and application examples. Finally, Chapter 6 contains discussion and conclusions regarding future research directions.

II. OPERATIONAL TRAFFIC-RESPONSIVE FREEWAY RAMP METERING ALGORITHMS

II.1 INTRODUCTION

The simplest form of ramp metering consists of a pre-timed operation, often termed open-loop control, where the ramp meter operates on pre-determined metering rates for a particular control period. The metering rates are generally based on observed past patterns of flow. Despite the simplicity and low cost of pre-timed control, its effectiveness is substantially reduced during incidents or when freeway traffic demand undergoes unexpected, rapid changes. This is especially true since traffic flow in most urban areas often undergoes rapid variations, and the need for control is higher under such conditions.

To increase the responsiveness of ramp metering to changing traffic conditions in real time, traffic-responsive metering schemes have been developed in the U.S. and other countries. Traffic-responsive control makes decisions in real time based on information from the freeway and communicates the corresponding control action to the controller, i.e., the on-ramp signal. While variations of traffic-responsive control schemes are numerous, operating metering systems can be categorized into two groups in terms of the way metering rates are determined in real time: automatic rate-selection systems and on-line rate-calculation systems. In automatic rate-selection, metering rates are selected from pre-determined rate-tables and thresholds using the information obtained from freeway loop detectors. By contrast, in on-line rate-calculation, metering rates are computed in real time following simple procedures. Each type of metering can be applied to either an isolated system or to an integrated one with multiple on/off ramps. This chapter reviews the freeway metering algorithms operating in major U.S. cities and other countries.

II.2 U.S. METERING SYSTEMS

II.2.1 Automatic Rate-Selection Systems

Detroit

The Detroit metering system employs a simple procedure with one upstream detector station measuring the occupancy of a center lane. One rate applies to the entire control period, and the upstream occupancy measurements are used to determine when meters should be turned on or off. The occupancy threshold at which meters are turned on ranges from 10 to 13 %, and the threshold at which they are turned off lies between 6 and 9 % depending on location (Kostyniuk, 1988).

Los Angeles and San Diego

Metering decisions are based on one mainline detector station located immediately upstream of the ramp to be metered (feedforward control). Selection of rates occurs every 6 seconds using the 1-minute lane-average volume or occupancy; the 1-minute average is updated every 6 seconds, and up to 6 lanes are monitored. Each ramp uses its own volume/occupancy thresholds and rate-table consisting of 15 rates. However, at each selection, the controller cannot impose a rate that differs more than one rate-step from the previous rate value. When traffic speed is higher than 50 mph, volume thresholds are used; occupancy thresholds are used otherwise (Stevens, 1990).

In San Diego, the occupancy measurements of a downstream detector station, whose location is pre-determined, are sometimes used (feedback control) in addition to the upstream information. In those cases, the rates corresponding to the upstream and downstream occupancies are compared and the more restrictive rate is selected (Raffat, 1990).

Milwaukee

Rate-selection occurs every 2 minutes using the smoothed measurements of volume and occupancy of a center lane, either upstream or downstream, depending on the geometric conditions of a given ramp. Most ramps use a common set of volume/occupancy thresholds and a common rate-table consisting of 6 rates (Warren, 1990).

Chicago

For each ramp in the Chicago metering system two detector stations are used, i.e., the adjacent upstream station and a downstream station whose location is pre-determined based on past experience. In both stations only the center-lane occupancy is measured. Rate-selections are made every 1-minute by comparing the 1-minute occupancy measurement of each station against the other, and by selecting the more restrictive rate using pre-determined thresholds. Most ramps use a common set of occupancy thresholds and a common rate-table consisting of 5 rates. Table 1 summarizes a typical set of occupancy-based metering rates from the current Chicago system (McDormott, 1979).

Table 1. Occupancy Thresholds (Chicago)

Mainline Lane Occupancy	Metering Rates
Less than 20%	12+ veh/min
20 - 22	10
22 - 25	8
25 - 27	6
27% or More	4

(Source: Chicago Area Expressway Surveillance and Control. McDermott, J., et al., 1979.)

Denver

The Denver metering system uses both local and coordinated control strategies depending on the existence of congestion. In local control, each ramp meter uses volume/occupancy thresholds to select one of 6 metering rates every 20 seconds. Mainline primary and secondary detectors immediately upstream of the ramp are used to determine the volume, speed and occupancy of each lane. The secondary detectors are used as a backup for the primary detectors. The measurements are smoothed to prevent rapid switching between rates. Queue detectors are installed near the entrance of the ramp to sense when vehicles are backing toward the cross street. When the occupancies of the queue detectors exceed a pre-determined threshold, the controller overrides the normally selected metering rate with less restrictive rates until the backup is reduced to an acceptable level.

If ramp A is "congested", i.e., the ramp is governed by either the most restrictive rate or by queue override, a system coordination plan is placed into operation. First, the ramp that is immediately upstream to ramp A is set at a rate which is one-step more restrictive than the rate determined by local control for that upstream ramp. If ramp A remains congested during the next 20 -second interval, the two ramps that are immediately upstream to ramp A are set at rates that are one-step lower than the rates determined by local control for the two ramps. The coordination continues by considering more upstream ramps, and stops at the time that ramp A is not congested (Lipp, 1991).

II.2.2 On-Line Calculation Systems

Seattle

The Seattle, Washington, metering system combines two methods for determining ramp rates. In particular, in addition to the automatic rate-selection procedure based on local occupancy measurements, it uses an on-line calculation procedure, called, "the bottleneck algorithm", to determine a rate every 20 seconds.

In the bottleneck algorithm, a freeway section is defined as a section between two consecutive loop detector stations on the freeway. The algorithm first determines if occupancy at the downstream detector station of a section exceeds a threshold occupancy i.e., the occupancy at which that section approaches capacity. If the threshold value, generally 18%, is exceeded, volume data are also checked to determine if more vehicles enter than leave that section. If vehicles are stored in a freeway section, the metering rates of the upstream on-ramps which influence that section are reduced to compensate for the stored vehicles. The reduction rate for each upstream ramp is calculated using weighting factors assigned to each ramp. The resulting metering rate of ramp i can be expressed as follows:

Bottleneck Rate for ramp i at time t =
[Rate at time t-1] - [Number of stored vehicles in a section * $W_i/\Sigma W_j$]

where, W_j, the weighting factor for ramp j, is determined from experience (Jacobson, 1989).

Richmond

The metering system in Richmond, Virginia, determines metering rates of a control section, consisting of up to 10 links, based on predicted link demand and capacity every 1 minute. A link normally begins at an exit ramp immediately upstream of an entrance ramp (Figure 2). To determine metering rates of a control section, the maximum permissible volume for each link, defined as the maximum volume that will not saturate the immediate downstream link, is first found. This volume is calculated from the capacity on that link and the capacity on the merging segment of the ramp immediately downstream of that link as follows:

$$RDL(n) = \min \{Q(n), F(n)C(n), [F(n+1)CC(n+1)-EQV(n+1)RL(n+1)]/P(n+1)\}$$

where, $RDL(n)$ = maximum permissible volume on link n,

$Q(n) = [RDL(n+1)-RL(n+1)]/P(n+1)$,

$P(n)$ = fraction of volume remaining on the mainline after vehicles have exited at the link n exit ramp,

$C(n)$ = current estimated link capacity on link n,

$F(n)$ = desired level of service coefficient on link n,

$CC(n)$ = current estimated merge capacity on link n,

$EQV(n)$ = ramp vehicle equivalency at the mainline-ramp merging area on link n,

$RL(n)$ = minimum metering rate on link n.

For each section, the above calculation starts with the furthest downstream link of that section and ends with the furthest upstream link. The metering rate of each link is, then, determined as the difference between maximum permissible volume and "predicted mainline-arrival demand" from the immediate upstream link. The predicted demand for link n is determined as follows:

$$PRD(n+1) = \max \{P1(n+1), P1(n)+RRVOL(n)-QE(n+1)\}$$

where, $PRD(n)$ = predicted arrival demand on link n,

$P1(n)$ = measured mainline volume on link n during previous time interval,

$RRVOL(n)$ = measured on-ramp volume on link n during previous time interval,

$QE(n)$ = predicted exit volume at the exit ramp on link n.

Currently historical data are used for QE(n), i.e., the exit volume in link n, with real-time adjustments. Finally, RC(n), the metering rate for link n is determined by

$$RC(n) = \min \{ [RDL(n)-PRD(n)], [F(n)CC(n)-PRD(n)]/EQV(n) \}$$

The metering rate calculation starts with the furthest upstream link and ends with the furthest downstream link in a section (Hall, 1990).

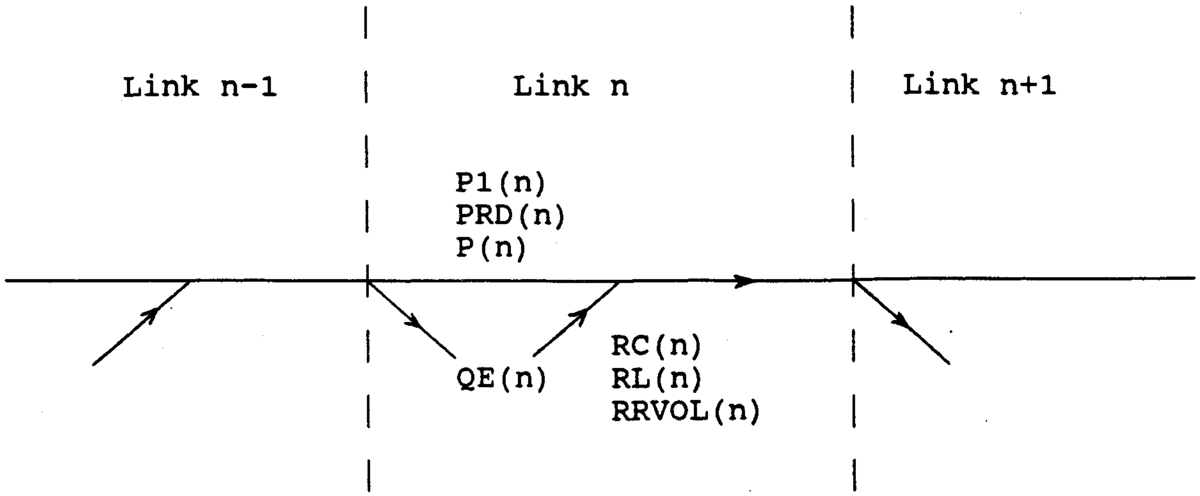

Figure 2. Link Configuration

New York: Long Island Expressway

Ramp metering is a subsystem of the Integrated Motorist Information System (IMIS), which is designed to control freeway corridors. The metering subsystem has the capability of both central and local modes of control for approximately 70 entrance ramps.

In the local control mode, the ramp controller uses the pre-determined cycle length selected on a time-of-day and day-of-week basis. This mode begins operating automatically if the central computer or communication fails. In the central control mode, metering rates of several entrance ramps are determined simultaneously based on predicted demand and downstream capacity. Demand is predicted at the entry points of each entrance ramp using a dynamic assignment approach for the whole freeway corridor (Zobe, 1982).

II.3 MINNEAPOLIS SYSTEM

In Minneapolis metering rates at a ramp are determined based on the lane-average volume data from a mainline station upstream, and the lane-average occupancy data from 5 mainline stations downstream of that ramp. For each ramp, traffic engineers have used historical data to derive a set of volume/occupancy thresholds and a rate-table consisting of 6 metering rates.

Rate selection, every 30 seconds, is based on two measurements, i.e., the 1-minute upstream volume and the highest 1-minute occupancy among 5 downstream detector stations. For instance, Figure 3 indicates that rate selection at the 98N entrance ramp is based on the 1-min upstream volume from station 39, and the highest 1-min occupancy amongst stations 40-44 downstream from the ramp. Using the pre-determined volume/occupancy thresholds and rate-table for the ramp, two rates are determined based on the two measurements, and the more restrictive rate is selected (Lau, 1990).

For each entrance ramp in the test site, Figures 4 and 5 summarize the volume/occupancy thresholds and the metering rates corresponding to volume and occupancy levels. As an example, from Figure 5, there are three occupancy thresholds at ramp 98N, corresponding to four metering rates. When the highest 1-min downstream occupancy is below 20%, the corresponding metering rate #3 is 500 veh/hr. For occupancies between 20% and 25%, the rate is 400 veh/hour; similarly, when occupancy is between 25% and 40%, the rate is 300 veh/hour. Finally, for higher occupancies rate

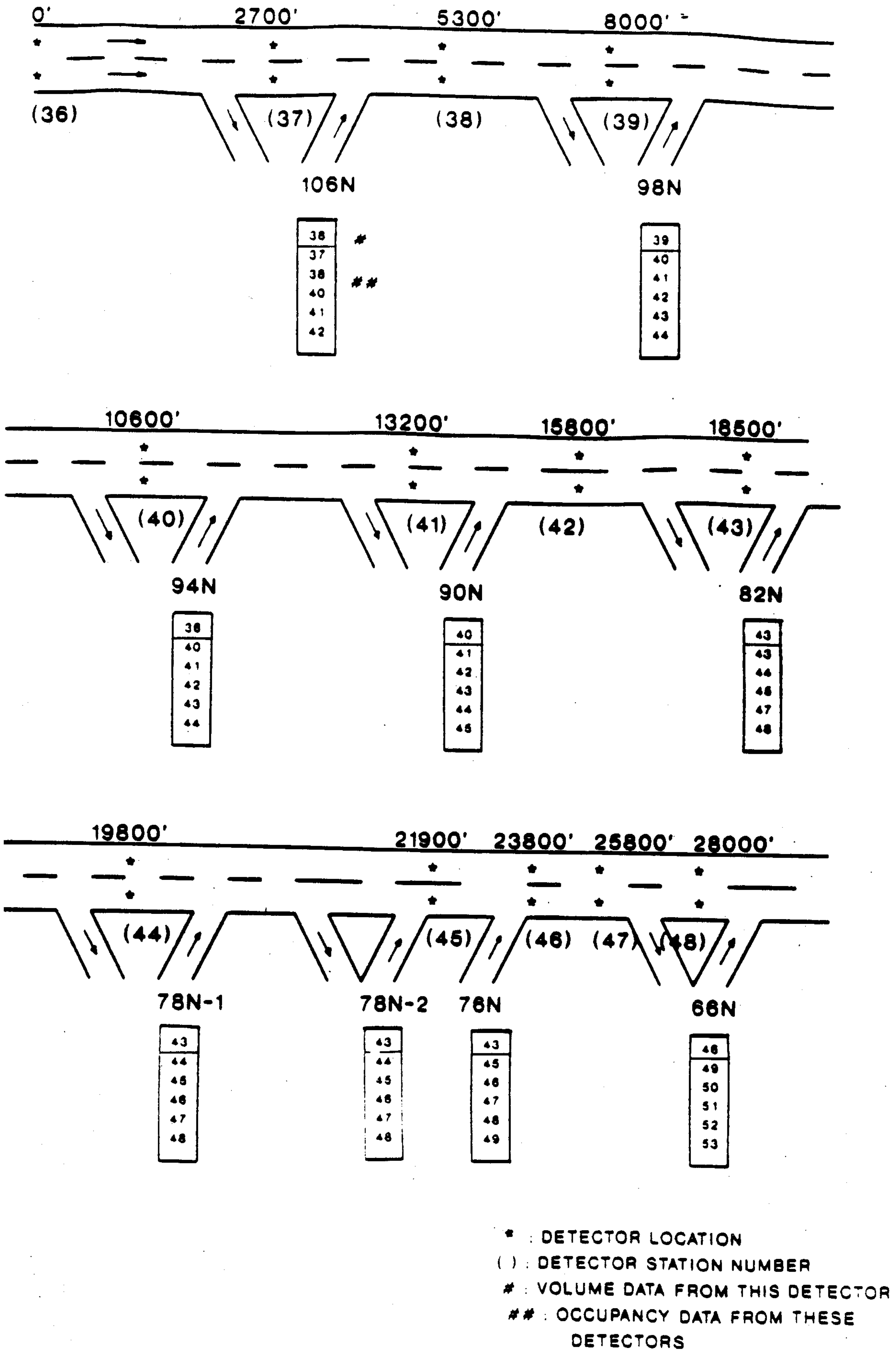

Figure 3. Freeway Test Section, I-35W, Minneapolis

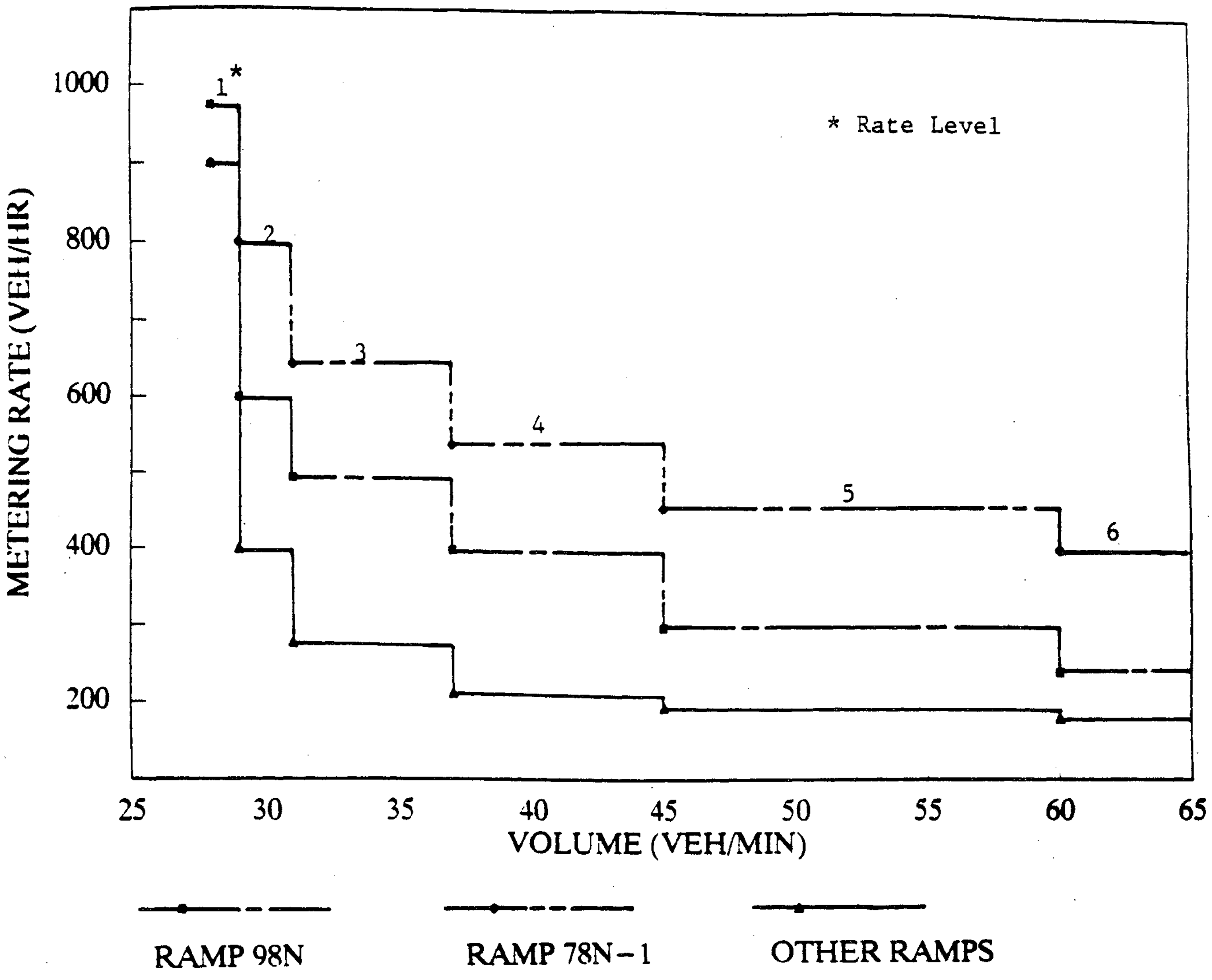

Figure 4. Metering Rate: Volume Thresholds

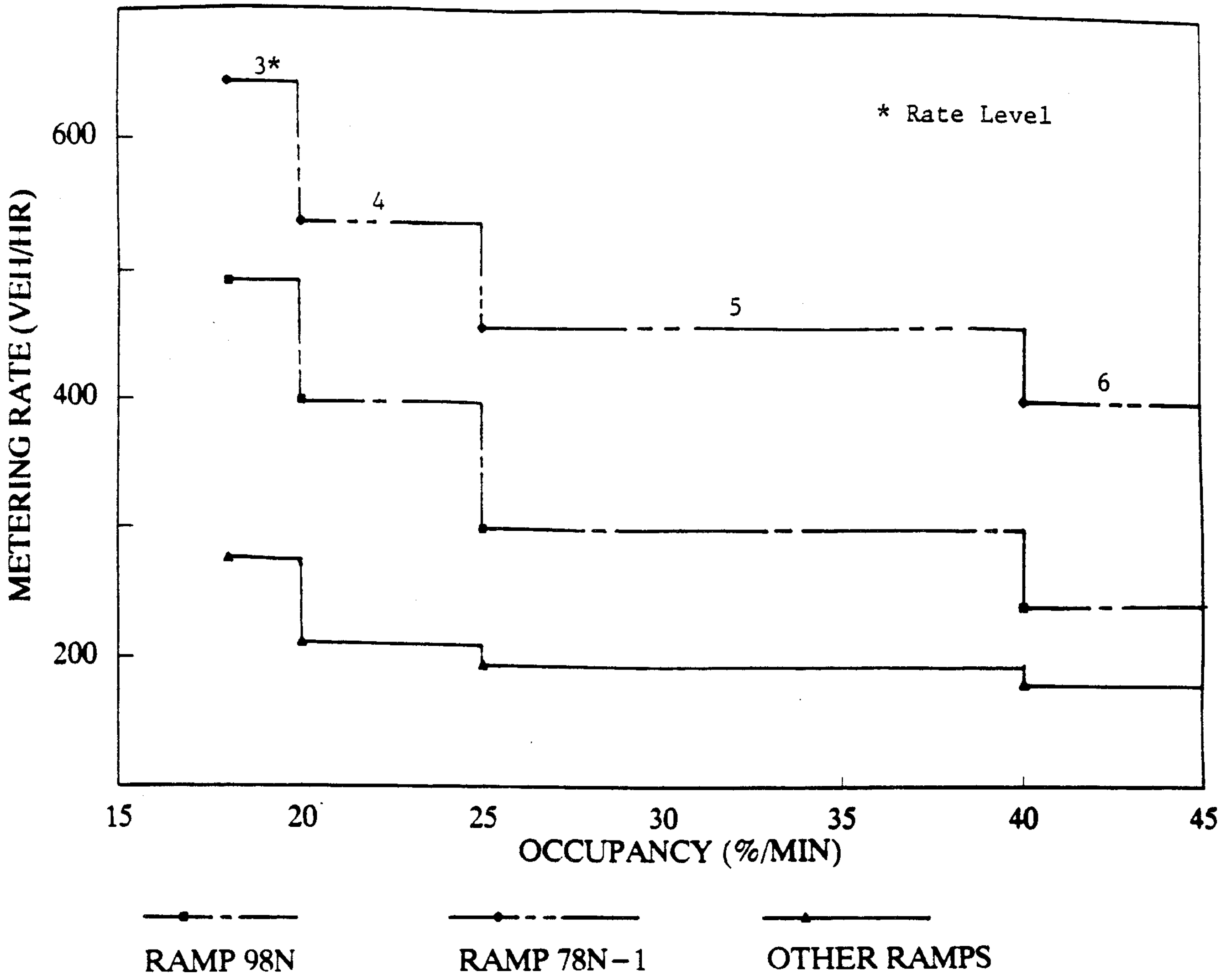

Figure 5. Metering Rate: Occupancy Thresholds

#6 would be required, corresponding to 250 veh/hour.

To determine the actual metering rate at the ramp, both Figures 4 and 5 are used. For instance, let the highest 1-min downstream occupancy at ramp 98N be 26 percent and the 1-min upstream volume at station 39 equal 62 veh/minute. Whereas the occupancy-based rate is #5 i.e. 300 veh/hour, as Fig. 5 suggests, the actual rate will be the more-restrictive #6 i.e. 250 veh/hour, which corresponds to the given volume as Fig. 4 indicates. Therefore, in this case, the control action is volume-based.

II.4 NON-U.S. SYSTEMS

II.4.1 Canada: Queen Elizabeth Way, Ontario

The Queen Elizabeth Way metering system in Mississauga, Ontario, employs an automatic rate-selection procedure with pre-determined rate-table and volume/occupancy thresholds. For each ramp, the rate-selection process takes place every 30 seconds using 1-minute average lane-volume/occupancy measurements of 3 detector stations assigned to that ramp. The 3 stations usually consist of one immediate-upstream (volume) and two downstream (occupancy) detector stations. All ramps share a common set of occupancy thresholds and a common rate-table consisting of 5 rates, but the volume thresholds for the upstream station vary from ramp to ramp. Table 2 illustrates the volume/occupancy thresholds and corresponding metering rates for the Ontario system.

Each ramp uses a set of backup stations from which it obtains data in the event that one or more of the downstream and upstream stations are not in operation or are transmitting inaccurate data. If the backup stations malfunction, the metering rates are determined from the 5-minute average measurements taken at the same time of the day, one week earlier (Fox, 1990).

II.4.2 Europe: Netherlands, Great Britain and France

In Europe, where urban freeways are not common, most ramp meters have been treated as traffic signals adopting the intersection timing hardware concept of cycle time, i.e., vehicles are allowed to enter the motorway continuously during the green signal of the meter. Most metering systems are still in an experimental stage.

Table 2. A Typical Set of Thresholds (Ontario System)

Rate Code	1	2	3	4	5
Local Station Occ 'LOCOCC'	20%	30%	40%	50%	70%
Downstream Station Occ 'DOWOCC'	20%	30%	40%	50%	70%
'CTSVOL' 'CTSSEC'					
Ups. st. vol.-Cont. Sect. 1	70	80	90	100	120
Ups. st. vol.-Cont. Sect. 2	70	80	90	100	120
Ups. st. vol.-Cont. Sect. 3	65	75	85	95	110
Ups. st. vol.-Cont. Sect. 4	65	75	85	95	110
Ups. st. vol.-Cont. Sect. 5	55	65	75	85	100
Ups. st. vol.-Cont. Sect. 6	55	65	75	85	100
Ups. st. vol.-Cont. Sect. 7	45	55	75	65	100
Ups. st. vol.-Cont. Sect. 8	45	55	75	65	100
Ups. st. vol.-Cont. Sect. 9	45	55	75	65	100
Ups. st. vol.-Cont. Sect. 10	70	80	75	90	110

(Source: Ramp Metering Algorithm Description. Fox, M., 1990.)

Netherlands: A-10 Motorway

The metering system in the motorway A10-west, near the Coentunnel in Amsterdam, adopts an on-line calculation procedure, that computes the metering rate every 30 seconds using the information from the detector located 1.5 km upstream of the ramp. A secondary backup detector station is also located upstream near the ramp. Both stations are equipped with double loop flow/speed detectors. The smoothed 30-second flow is calculated from the following formula:

$$\text{New Flow} = \alpha * \text{Current Flow} + (1 - \alpha) * \text{Old Flow}$$

The cycle time of the entrance ramp signal is calculated from

$$\text{Cycle time} = 3600 / (\text{Capacity} - \text{New Flow}).$$

When speed on the motorway drops to a value lower than 35 km/hr, the cycle time is set to its maximum value. After speed reaches a value higher than 50 km/hr, cycle time is again computed from the above formula (Buijn, 1990).

Great Britain: M6-Motorway

The on-line M6-motorway metering system determines the state of an entrance -ramp signal, i.e., red or green, every second based on the upstream and downstream flow-speed measurements. Each time, the downstream capacity (defined as the likely maximum flow at current driving conditions) is estimated by comparing current-day flow/speed measurements with historical data from a table. Traffic demand is also measured on the main motorway and the entrance ramp every second. By comparing the measured mainline and entrance ramp demand with the estimated downstream capacity every second, the metering algorithm determines whether the signal should remain green or change to red (Keen, 1986).

France: Boulevard Peripherique

The Boulevard Peripherique system uses one detector station that measures occupancy downstream of the merge area. The metering rate is calculated every 40 seconds from the following formula:

Current Rate = Previous Rate + K [Desired Occupancy - Current Occupancy]

where K is a parameter calibrated off line. For instance, experiments indicated K = 70 vph and a detector location 40m downstream of the beginning of the entrance ramp to be adequate. However, these parameters would have to be recalibrated at different locations and ramp types (Hadj-Salem, 1990).

II.4.3 Japan: Hansin Expressway, Osaka

The traffic control system of the 81-km Hansin Expressway, covering the Osaka-Kobe area, includes two types of control, each of which is applicable to a specific state of traffic flow. The two types of control are linear programming control and sequential ramp closure.

In particular, for stationary traffic flow, where there is no traffic jam and no sudden changes of traffic volume entering or leaving the expressway system, the volume of traffic entering every entrance ramp is determined through linear programming control. Such control attempts to find optimal rates that maximize the total number of vehicles entering the expressway without causing any traffic jam in any section of the roadway. The origin-destination data needed for linear programming are derived off-line using the entropy method (Sasaki, 1968).

When mainline flow and entering traffic volume vary widely over a short time interval and congestion is expected to occur soon thereafter, sequential control is implemented by closing the upstream entrance ramps. Criteria for determining which ramps should be closed, and in what order, are derived off-line through simulation (Kometani, 1974).

II.5 SUMMARY

As reviewed in this chapter, the most common problem in existing operational, traffic-responsive, ramp metering systems is that they still cannot reflect network-wide traffic conditions and flow interactions between freeways. To be sure, most operational traffic-responsive metering systems set their metering rates based on local congestion levels usually represented by occupancy measurements. While a few systems attempt to reflect system-wide traffic conditions by making adjustments that are based on downstream conditions, existing metering algorithms tend to be reactive rather than predictive. More specifically, they use previously collected data, generally 1-minute old, for calculating metering rates, thereby reacting to freeway conditions rather than preventing congestion before it happens. Further, in the automatic rate-selection systems, the updating procedures for thresholds and rate-tables are not well defined and mostly depend on past experience.

Capacity estimation is an additional important issue. While most on-line rate-calculation procedures require estimation of capacity of the freeway section to be controlled, they use predetermined, constant capacity values during the entire control period without considering the possible influence of the changing traffic environment on capacity.

Finally, no on-line optimization procedure based on realistic traffic flow models has been implemented in real life applications. While the metering system in Osaka, Japan, employs an optimization approach, the over-simplification associated with the linear programming restricts the effectiveness of system-wide control.

To improve traffic management in freeways, enhancement of existing metering strategies as well as development of efficient, realistic optimal control strategies is of critical importance. The following research directions have been identified to meet both short- and long-term needs for improving freeway management in real time.

1) Development of an efficient procedure for updating thresholds in automatic rate-selection metering systems.
2) Development of predictive algorithms that can be used in combination with existing automatic rate-selection systems.
3) Development of an efficient optimal control structure, that is computationally feasible and leads to robust control solutions, for real-time management of small freeway networks.
4) Development of an efficient on-line optimization algorithm that can be implemented for real time management of large freeway networks using parallel processing techniques.

The current project will address the first research direction by developing a control-emulation method for evaluating existing automatic rate-selection strategies.

III. ADVANCED CONCEPTS FOR OPTIMAL FREEWAY CONTROL: HIERARCHICAL CONTROL

III.1 INTRODUCTION

The most advanced concepts in freeway control employ hierarchical control structures, where overall freeway control is decomposed into several components or layers, such as prediction, optimization and direct control in Figure 6. The main objective of hierarchical control is to achieve computational feasibility and robustness of control solutions by determining the system-wide, nominal metering rates first and then adjusting them according to real traffic conditions. Previous efforts have employed the hierarchical control concept based on two, three or four layers.

III.2 OVERVIEW OF HIERARCHICAL CONTROL STRATEGIES

III.2.1 Two Layer Control

In two layer control, the first layer determines nominal metering rates and the second layer adjusts the nominal metering rates depending on the traffic conditions. Yuan and Kreer (1968) first proposed a two layer hierarchical control including a steady state component and a time varying component. In the steady state component, a linear programming problem is formulated from a simplified traffic continuity equation, and is solved for a nominal steady state condition including ramp metering rates and freeway section densities. The time varying component attempts to adjust the nominal ramp metering rates by minimizing the deviations from the steady state condition, i.e. minimizing the square of the difference between measured densities and steady state densities.

A similar two layer control system, that includes a nominal control layer and a traffic-responsive control layer, has been proposed by Isaksen and Payne (1973). The nominal control layer adopts a time-of-day metering rate as the nominal metering rate and, then, simulates the freeway system with this nominal control to obtain nominal states, densities and speeds of each freeway section. The direct control layer attempts to return the freeway to nominal conditions after a disturbance using an augmentation method to decompose and solve a linear regulator which minimizes the deviation from the nominal conditions.

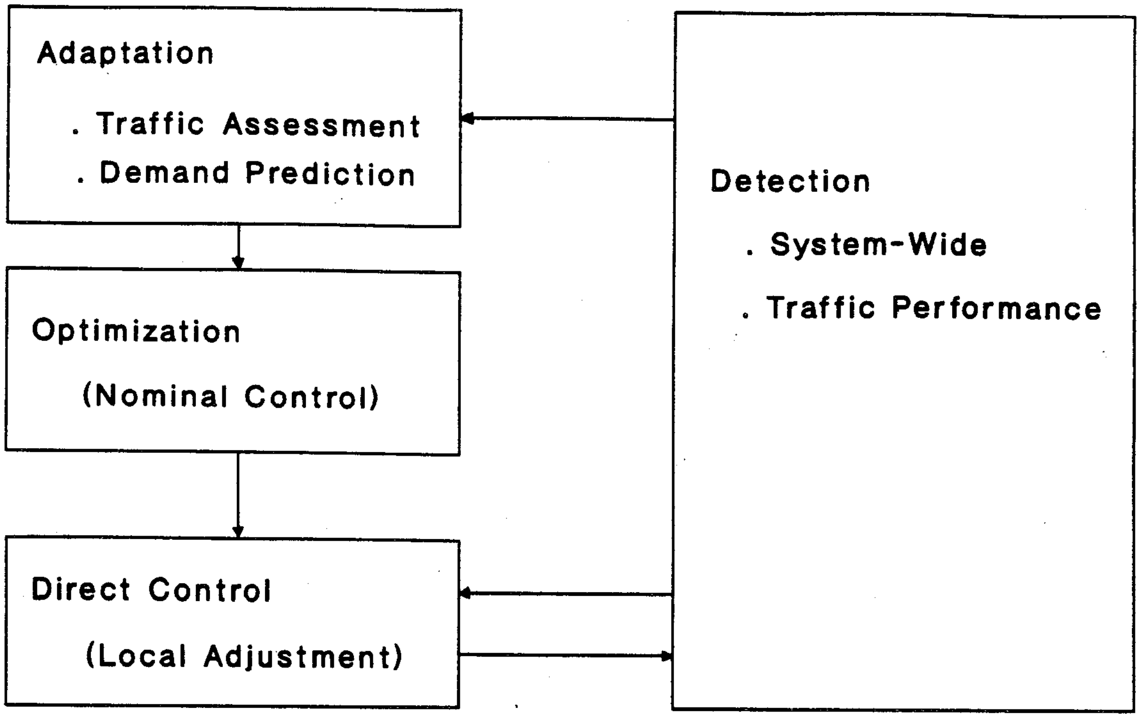

Figure 6 Hierarchical Traffic Control System

III.2.2 Three Layer Control

Papageorgiou (1983) proposed a three layer control system consisting of an optimization layer, a direct control layer, and an adaptation layer. The optimization layer obtains an optimal nominal condition by formulating and solving a linear programming problem similar to Wattleworth and Berry's model developed in 1965. The direct control layer minimizes the deviation from the nominal condition to maintain the nominal states. The adaptation layer predicts parameters resulting from slow traffic disturbances to initialize a new optimization. Payne and his colleague (1985) proposed a similar control system including an allocation component that determines a set of nominal ramp metering rates, a surveillance component that estimates traffic states from measured data, and a regulation component comparing the estimated traffic states to expected traffic states in order to generate adjustments to the nominal ramp metering rates.

III.2.3 Four Layer Control

Four layer hierarchical control, as proposed by May (1976), includes initialization, estimation, optimization, and tactics processors. Initialization makes the decision about whether to use control or not depending on the current traffic conditions. The estimation processor predicts parameters, e.g., demand and capacity, for optimization. Optimization formulates and solves a linear programming problem to obtain optimal metering rates. The tactics processor compares the predicted data to the measured data in order to determine whether to go back to optimization or to continue operating at current metering rates.

Table 3 summarizes the control concepts proposed for hierarchical ramp metering control. In addition to predicting parameters and deciding whether to use control, the actual tasks for determining metering rates focus on two layers, optimization and direct control. The optimization layer determines nominal conditions including nominal metering rates and nominal states according to the predicted traffic conditions. The direct control layer adjusts the nominal metering rates to reduce the influence of traffic disturbances and return the freeway to the nominal conditions. These two layers will be discussed further in sections III.2 and III.3, respectively.

Table 3: Hierarchical Control Structures

Control Frame	Features	Research Team
Two Layer	1) Steady State Component 2) Time Varying Component	Yuan & Kreer, 1968
	1) Nominal Control 2) Traffic Responsive Control	Isaksen & Payne, 1973
Three Layer	1) Optimization Layer 2) Direct Control Layer 3) Adaptation Layer	Papageorgiou, 1983
	1) Allocation 2) Regulation 3) Surveillance	Payne et al., 1985
Four Layer	1) Initialization Processor 2) Estimation Processor 3) Optimization Processor 4) Tactics Processor	May, 1976

III.3 OPTIMIZATION STRATEGIES AND TRAFFIC FLOW MODELS

Traffic flow models employed in the optimization layer fall into three categories, namely input-output, continuity, and high order models. Input-output models calculate traffic flows based on the information from the origin-destination matrix and upstream demand. Continuity models estimate traffic densities according to the traffic conservation principle. High order models employ a continuity model and a momentum equation to estimate the dynamic changes of traffic density and vehicle speed.

III.3.1 Input-Output Model

Employing an input-output traffic flow model to estimate freeway traffic flow as well as formulating and solving a linear programming problem to obtain optimal ramp metering rates were first proposed by Wattleworth and Berry (1965). According to the information of a pre-determined ramp origin-destination (O/D) trip matrix, the traffic volume on a freeway section is calculated by summarizing the upstream demand that crosses this section. More specifically, a static traffic state is estimated depending on the information from entrance-ramp volume (input) and exit-ramp volume (output). Based on this input-output traffic flow model, a linear programming problem is formulated to maximize ramp entrance volumes subject to the capacity of each freeway section, and is solved for the optimal ramp metering rates.

In this model, the accuracy of the O/D information influences the optimal solution. The time duration of the optimal control operation and the length of the freeway system must be selected to satisfy the assumption that steady state traffic flow is fully developed in the system during the period. Since the real traffic volume fluctuates and can be higher than the volume estimated from the O/D matrix, the capacity of each section, that constrains the section volume, must be adjusted to be less than the real section capacity so that the real freeway system using this optimal control will not break down when a small traffic disturbance occurs. In addition, traffic congestion conditions, such as queue spillback, cannot be described by this input-output model which does not consider time variation.

A similar model (Wang and May, 1973; Ovaici et al., 1975; Imada and May, 1985), FREQ, was extended to, 1) optimize different objective functions, such as maximizing total passenger-miles of travel, with the aid of classifying ramp entering vehicles by occupancy, 2) handle traffic variations using capacity and metering buffers, and 3) manage traffic congestion by adding freeway queue storage in each time slice and passing it to the next time slice.

Papageorgiou (1980) considered the simple time evolution of freeway flow according to the delay between a volume change at a ramp and its subsequent disturbance at a freeway point downstream. The same objective function as Wattleworth and Berry's (1965) subject to more constraints, such as freeway section capacities and the queue length at entrance ramps, is formulated as a linear programming problem. In the problem, the capacity of a freeway section is reduced from its maximum flow rate by a small amount called capacity buffer such that, in the real system, the flow perturbation around the optimal solution is stable, i.e. the flow rate does not exceed the maximum flow rate. In addition, since the model considers time evolution, the optimal solution is closer to the actual, fluctuating traffic volume, than Imada and May's solution (1985). A relatively small capacity buffer is enough to handle traffic variations.

In general, input-output models are linear and static. They do not consider dynamic effects which can be more accurately described by non-linear models. Since frequent state transitions and rapid flow variations are observed in most real systems, the control strategies determined through steady state traffic conditions are of very limited use even if they are updated periodically.

III.3.2 Continuity Model

Based on the vehicle conservation principle, the continuity equation has been employed to describe the dynamics of freeway flow (Yuan and Kreer, 1968; 1971; Kaya, 1971). In these models, the freeway system is separated into several sections. Traffic flow in each section is represented by its density that is considered as a state variable. The continuity equation is employed as the state equation to estimate the changes in the state variables. Through the state equation, traffic flow of each section can be estimated for every time increment and is governed by the control variable, i.e. the ramp metering rate.

Yuan and Kreer (1968, 1971) simplified the continuity equation so that the flow rate of each freeway section can be formulated as a linear combination of upstream mainline demand and ramp metering rates. Using this simplified continuity model, they formulated a linear programming problem that maximizes weighted ramp demand and queue length, i.e. allowing high entering volume and balancing queue length, subject to capacity of each freeway section (Yuan and Kreer, 1968). They also considered a quadratic objective function formed by the square of the difference between ramp metering rate and equivalent demand flow rate. They solved this quadratic programming problem using Frank and Wolfe's method for nominal conditions which include metering rates and section densities (Yuan and Kreer, 1971). Since the continuity model was linearized, the simplified model is similar to the input-output model proposed by Papageorgiou (1980) and, thus, loses the ability of describing traffic flow dynamics.

Kaya (1971) proposed a non-linear optimization problem that minimizes the total travel time including waiting time, merging time, and travel time within the freeway system subject to the original non-linear continuity model for determining an optimal nominal metering rate. Using a polygonal approximation technique, the non-linear optimization problem was transformed into a linear programming problem with more constraints and variables. Even though the new problem is linear and has more constraints and variables, it can be solved easily and can still estimate dynamic traffic flow. Depending on the ways in which the dynamic continuity model is formulated and linearized, it could precisely estimate freeway section densities by the hour, minute, and even second. However, using the continuity model requires a known speed-density or volume-density relationship for computing speed and volume.

Stephanedes and Chang (1991) proposed an optimal ramp-metering control model for freeway corridors. In the model, non-linear continuity equations were developed for the major components of a freeway corridor as state equations. The objective function, total travel time of the freeway corridor, was formulated as a linear function of state variables. The optimal control problem was converted into an optimization problem without constraints using Lagrange multipliers, and solved by the Conjugate Gradient method for optimal ramp-metering rates. As in Kaya's work, this model can estimate dynamic corridor traffic flow, but requires a known speed-density relationship. In addition, whereas that work was limited to the freeway, this model has been designed for application that include the freeway, adjacent arterials, and other major components of freeway corridors.

III.3.3 High-Order Traffic Flow Model

A more complex model for describing traffic dynamic behavior was developed by Payne (1971). In addition to traffic flow conservation, the model considers the dynamic changes of speed on the freeway using a momentum equation adapted from fluid mechanics.

A similar traffic flow model was developed by Blinkin (1976). Additional functions, such as lane changing and speed control, are formulated in the continuity and momentum equations to estimate dynamically interactive traffic flow behavior between lanes and analyze the influences of speed-limit control. In his optimization problem, density, speed, and queue length are the state variables, and the parameters for ramp entrance control, speed control and lane changing control are the control variables. The optimization objective seeks to minimize the oscillations of the state variables, i.e. minimize the square of the difference of the state variable values at adjacent time intervals, subject to a multi-function dynamic traffic flow model and the queue length at each ramp.

The solution to this problem formulation would require a non-linear programming method. Since the high-order model can describe detailed traffic dynamics, the resulting optimal metering rates could respond to non-recurrent traffic congestion on the freeway. However, since the optimization problem is more complex and has a larger number of state and control variables than the problem using a continuity model with the same space and time discretization, the solution would be more computationally intensive. Moreover, when the nominal states, density and speed, are delivered to a direct control layer as references for adjusting nominal control, the effort for real-time measurements of these states and real-time computations to determine the solution in the direct control layer is increased.

Cremer (1976) proposed an optimization problem formulation using a high-order traffic flow model for the state equations. In this model, two continuity equations are employed for each freeway section. One equation describes the traffic behavior of trucks, and the other models all vehicles. The objective of the problem is to maximize the summation of the product of speed and volume as well as a weighted reciprocal of the metering rate over time and space. This objective implies that the freeway can operate with a high volume at high speed, and can maintain ramp entering volumes at high levels. A parameter optimization technique was used to solve the problem.

Table 4 summarizes the major traffic flow models that have been proposed for the optimization layer. Since the high-order formulation can estimate dynamic traffic behavior and performance in detail, high-order modeling in the optimization layer could lead to the required nominal conditions and could deal with non-recurrent traffic congestion. However, complex high-order models are difficult to solve and require a long computation time compared to input-output models and continuity models with the same discretization. In addition, the large number of nominal states increases the effort for measuring data and computing adjustments of nominal control in the direct control layer.

III.4 DIRECT CONTROL STRATEGIES

Returning the freeway to its nominal state by adjusting the nominal metering rates is the goal of the direct control layer. The major models that have been proposed for the direct control layer are summarized in Table 5. Most researchers employ a regulator to deal with direct control. A regulator adjusts the nominal metering rate by minimizing the deviation of measured data from the nominal conditions.

III.4.1 Regulator

Yuan and Kreer (1968) proposed a standard linear regulator as the time varying component in a two-layer control system for adjusting the nominal metering rates. They used a perturbation method to formulate the problem. In particular, the differences between the measured freeway section densities and the nominal densities as well as the differences between the executed metering rates and the nominal metering rates are defined as error terms. The regulator problem is formulated as the minimization of a linear quadratic performance function composed of the square of the error terms subject to the changes of the density error that are estimated by a linear combination of the density and metering-rate errors. The optimal solution of the regulator is the metering-rate error which is added to the nominal metering rate such that, following a disturbance, the adjusted metering rates

Table 4: Traffic Flow Models in the Optimization Layer

Traffic Flow Model	Features	Research Team
Input-Output Model	▪ Maximize Ramp Entering Volume ▪ Subject to Section Capacity ▪ Solved by Linear Programming ▪ Use Ramp Metering Rate as the Control Variable	Wattleworth & Berry, 1965 Wang & May, 1973 Ovaici et al., 1975 Papageorgiou, 1980 Imada & May, 1985
Continuity Model	▪ Maximize Ramp Entering Volume or Minimize Travel Time ▪ Subject to Continuity Equation ▪ Solved by Linear or non-Linear Programming Methods ▪ Use Ramp Metering Rate as the Control Variable ▪ Use Road Section Density as the State Variable	Yuan & Kreer, 1968; 1971 Kaya, 1970 Stephanedes & Chang, 1991
High- Order Model	▪ Maximize or Minimize Performance Index ▪ Subject to Continuity Equation and High-Order Speed Model ▪ Solved by non-Linear Programming Methods ▪ Use Ramp Metering Rate and other Variables as the Control Variable ▪ Use Road Section Density and Speed as the State Variable	Blinkin, 1976 Cremer, 1976

Table 5: Control Structures in the Direct Control Layer

Control Structure	Features	Research Team
Regulator	▪ Minimize Deviation from System Nominal Condition ▪ Solved by Centralized or Decentralized Methods	Yuan & Kreer, 1968 Isaksen & Payne, 1973 Payne et al., 1973 Looze et al., 1978 Papageorgiou, 1981 Goldstein & Kumer, 1982 Papageorgiou & Mayr, 1982 Payne et al., 1985
Local Feedback Control	▪ Minimize Deviation from the Downstream Nominal Condition	Hadj-Salem et al., 1990

would drive the freeway back to its nominal condition. The regulator problem is solved by the well known Riccati equation. However, as the system size increases, so does the computation time required for solving the Riccati equation; for large systems this may be impractical. In addition, using a linear function to estimate the changes of density error may not return the freeway to the nominal conditions if a large disturbance occurs.

A similar regulator problem including additional speed-error terms was decomposed into sub-systems, and solved by an augmentation method (Isaksen and Payne, 1973; Payne et al., 1973; Payne et al., 1985). The augmentation method, first, partitions the original regulator system into overlapping sub-systems by deleting weak coupling terms between sub-systems, such as far away downstream and/or upstream freeway section densities and speeds, so that the optimal solution of low-order sub-systems can be found. Then, these optimal controls are reassembled to construct a sub-optimal control for the original problem. Thus, the augmentation method can retain the coupling between sub-systems, and lead to a sub-optimal solution. In addition, since the original problem is formulated as a linear system, the controlled freeway is difficult to recover from a large traffic disturbance.

Looze et al. (1978) proposed a linear quadratic Gaussian (LQG) regulator for direct control. The control and state estimation in the original LQG regulator are decentralized into local sub-systems. The local control is formulated as a linear function of local state estimates to minimize a system-wide performance. The dynamic state estimator using a filter technique is composed of local state estimates, local measurements, and certain estimates which communicate with other sub-systems. The decomposed problem is converted into an equivalent parameter optimization problem. The new problem can be solved off-line by a modified Davidon-Fletcher-Powell algorithm to obtain a numerical solution to the static optimization. In the decomposed regulator, local control and the state estimator enhance the ability to adjust local disturbances and retain the coupling between sub-systems. However, in this case, only a sub-optimal solution is obtained. In addition, the stability problem is not assured in this complex system.

A multilevel decentralized control scheme, the cascading technique, was proposed by Isaksen and Panye (1973) (Goldsten and Kumar, 1982) to solve a linear quadratic problem. The cascading technique seeks to enlarge the original system in a systematic way, and partition the enlarged system into controlled sub-systems. The coupling between two sub-systems includes more then one section so that controlling a section could be handled by the adjacent controllers when the controller in that section breaks down. Therefore, this control system is more reliable than the control systems without the control coupling.

Papageorgiou (1981) proposed a non-linear quadratic problem that adopts the same linear quadratic performance function as Isaksen and Payne's model (1973) subject to non-linear state equations developed by Payne (1971). The problem was solved using decomposition methods including interaction prediction principle, interaction balance principle, penalty function method, and co-state coordination method (Papageorgiou and Mayr, 1982). Non-linear systems could manage large disturbances, but need more computation time to obtain the solution.

In summary, the computation time required of a large-scale regulator is high. Decomposing the regulator by deleting weak coupling between sub-systems could reduce the computation time, but can only lead to a sub-optimal solution. Linear regulators can be solved faster than non-linear regulators, but they may not draw the freeway back to its nominal state after a large disturbance. Moreover, stability and reliability problems must be considered prior to implementing such a regulator.

III.4.2 Local Feedback Control

A local feedback control system is a kind of regulator that controls an individual ramp with downstream measured data, and does not have any communication between ramps. It can be used in the direct control layer to keep the traffic of the ramp merging area around the optimal solution. Such control was proposed by Hadj-Salem, et. al., (1990) but is limited to comparing only one state measurement, downstream occupancy, to nominal occupancy. Since the feedback mechanism does not consider any coupling between sub-systems, the metering rate for one ramp could be obtained very fast by solving a small local optimization problem independently. However, maintaining the optimal system performance that results from the optimization layer would be difficult.

III.4.3 Local Feedforward Control

Local feedforward control is similar, in terms of performance, to local feedback control. The major difference between the two is in the type of data they use for adjusting metering rates. In particular, feedforward control uses upstream traffic data whereas feedback control uses downstream data. The simplest type of feedforward control determines the metering rate by substracting upstream volume measurements from the pre-defined capacity of the freeway. Feedforward control has been used in several freeway systems, such as Los Angeles and the A-10 motorway in the Netherlands, both discussed in Chapter II. Feedforward control can be easily implemented but, as in the case of feedback control, will have difficulty in maintaining the system performance at an optimal level.

III.5 SUMMARY

This chapter reviews hierarchical control strategies found in the literature for optimal freeway control. The simplest type of such control is performed in one layer. In particular, selecting or calculating metering rates based on real-time measurements could reduce the effects of traffic variations and congestion. However, this type of control cannot always make the necessary adjustments to the local metering rate and/or does not optimize the traffic performance of the freeway system. On the contrary, multi-layer hierarchical control optimizes system performance to determine a nominal metering rate (optimization layer) and adjusts this nominal rate based on current traffic conditions (direct control layer).

In the optimization layer, the control can be determined based on any of three major alternative formulations. An input-output traffic flow model would be the simplest option. Although such formulation would be easy to implement and would have low computation time requirements, it has substantial disadvantages. In particular, it requires origin-destination information, it is optimized under steady-state conditions, and cannot respond to heavy traffic congestion. Continuity traffic flow models can also be used to determine the nominal conditions of the control in the optimization layer. Such a modeling framework can optimize dynamic traffic performance, can estimate nominal densities, and can handle non-recurrent congestion. Compared to the input-output model, it requires more computation time and requires demand and diversion prediction information instead of O/D information. A more complex high-order traffic flow model which provides the capability of estimating dynamic traffic performance, density, and speed is adopted by certain researchers for determining nominal control in the optimization layer. The control formulation based on the high-order model can deal with non-recurrent congestion and would lead to more accurate estimates than the estimates resulting from a continuity model with the same discretization. Nevertheless, it requires more computation time and has a larger number of variables than that of the continuity formulation with the same space and time discretization.

In the direct control layer, the nominal control is adjusted using a regulator or a local feedback or feedforward control. Since rapid and large traffic flow variations are observed in most real systems, the direct control layer requires a large number of fast computations and solutions that are stable and robust. To simplify the problem and reduce the computation time, a large-scale regulator problem may be decomposed by deleting weak coupling between sub-systems. A linear regulator formulation can be solved faster than a non-linear one, but the adjusted control may not return the freeway back to the nominal conditions after a large disturbance. Although local feedback control may reduce the computation time required for adjusting the nominal control, it would not be able to maintain the system-wide optimal performance that results from the optimization layer.

IV. DEVELOPMENT OF A CONTROL-EMULATION METHOD FOR EVALUATING AUTOMATIC RATE-SELECTION STRATEGIES

IV.1 INTRODUCTION

The control-emulation method developed by Stephanedes, et al., (1992) for this study emulates real-time rate-selection metering using simulated freeway traffic performance when the freeway demand pattern and geometrics are known. Figure 7 illustrates the overall structure of the method, which consists of two major modules, i.e., ramp control and freeway performance modules, interacting continuously through time. The performance module determines traffic conditions at predefined detector locations and provides traffic data, e.g., volume and occupancy, to the ramp control module, which in turn selects the metering rates for the next time interval emulating a user-specified decision rule with the data provided by the performance module. The metering rates are then given to the performance module after a time delay, specified by the user depending on the system characteristics. Based on the metering rates and a known demand pattern, the performance module simulates the traffic performance until the next rate-selection time. The on-line interaction between ramp control and freeway performance continues for the whole control period, such as one or two peak hours. The performance module adopts a dynamic simulation model, which can determine the effects of frequent metering changes on traffic performance at any location along the freeway through time.

An important feature of the control emulation method is that it can reflect the characteristics of the communications system associated with a metering operation. In particular, it can explicitly treat the effects of time delays, i.e., the time needed to implement new metering rates after the rates are selected, on traffic performance. Various rate-selection time intervals, ranging from 2 seconds to a few minutes, can also be evaluated. The rest of this chapter describes the method and an example application in detail.

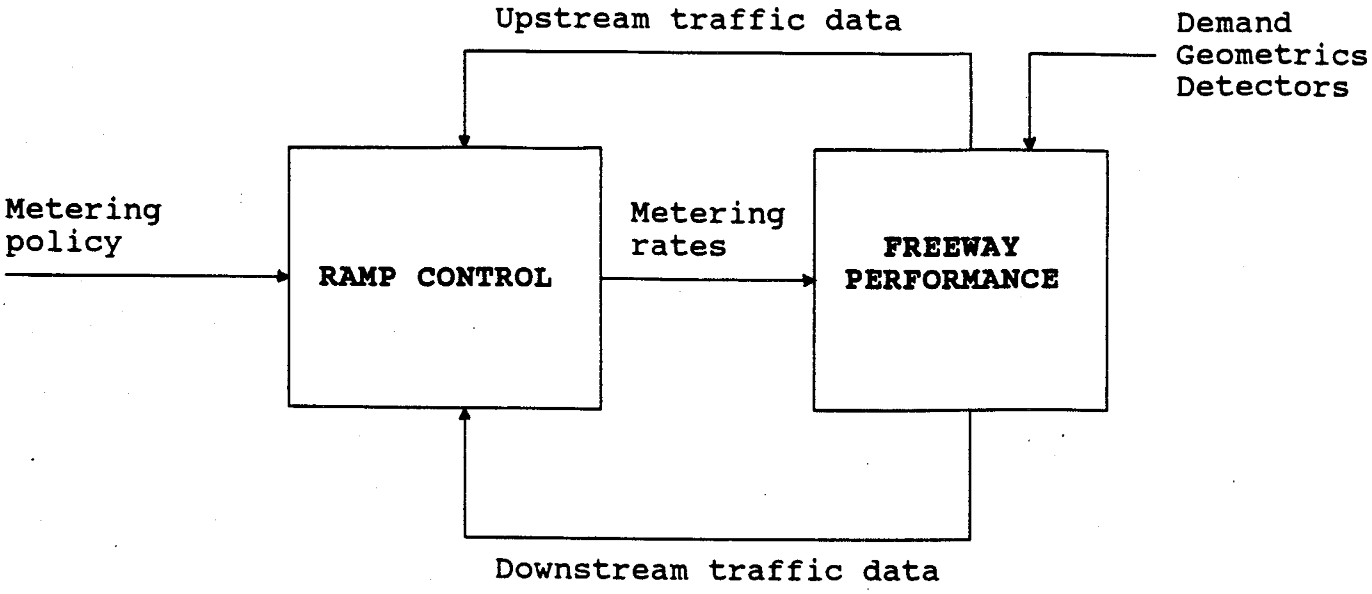

Figure 7 Control Emulation Structure

IV.2 RAMP CONTROL MODULE

The ramp control module emulates the on-line automatic rate-selection process and determines metering rates using the traffic information provided by the freeway performance module. From Figure 8, the key elements of the process include a set of volume/occupancy thresholds, a rate table and the locations of detector stations associated with each ramp. First, traffic volume/occupancy information at pre-defined detector locations are determined by the freeway performance module, which adopts a dynamic simulation methodology based on continuum modeling. Upstream volume information and the highest downstream occupancy among detector stations are then compared with the pre-determined volume/occupancy thresholds. Using the rate table, two rates are identified and the more restrictive rate is selected. Following a time delay, the freeway performance module uses that rate for determining the number of vehicles entering the freeway system.

The above metering policy can be further described using the volume-occupancy diagram. To illustrate, consider a typical volume/occupancy threshold policy used for all ramps except ramps 82N and 76N in the test section shown in Figure 3. The policy is described by curve AA' in the volume-occupancy space, Figure 9, where the whole space is subdivided into 4 regions, i.e., no-control, occupancy-dominant, volume-dominant, and common volume-occupancy (V-O) control regions. If a volume or occupancy measurement from the freeway detector stations falls into either the volume- or the occupancy-dominant region, the metering rate is determined by volume or occupancy respectively. In the common V-O control region the resulting metering rate is the same, whether it is determined by volume or occupancy. In Figure 4, the six metering rate levels are indicated by number, ranging from 1 to 6, where rate 6 is the most restrictive. For instance, if upstream volume is 40 veh/min and highest downstream occupancy is 30% per minute, control will be occupancy-based, and rate #5 will be activated at the ramp meter. If occupancy falls to 18% and volume remains the same, control will activate volume-based rate #4, a less restrictive rate.

In summary, the characteristics of each metering policy can be represented by the location and shape of a threshold curve in the volume-occupancy space. In the case of the Minneapolis traffic operation, any threshold policy represented by a curve below AA' in Figure 9 implies a stronger emphasis on volume control, whereas any curve above AA' denotes more occupancy-oriented control.

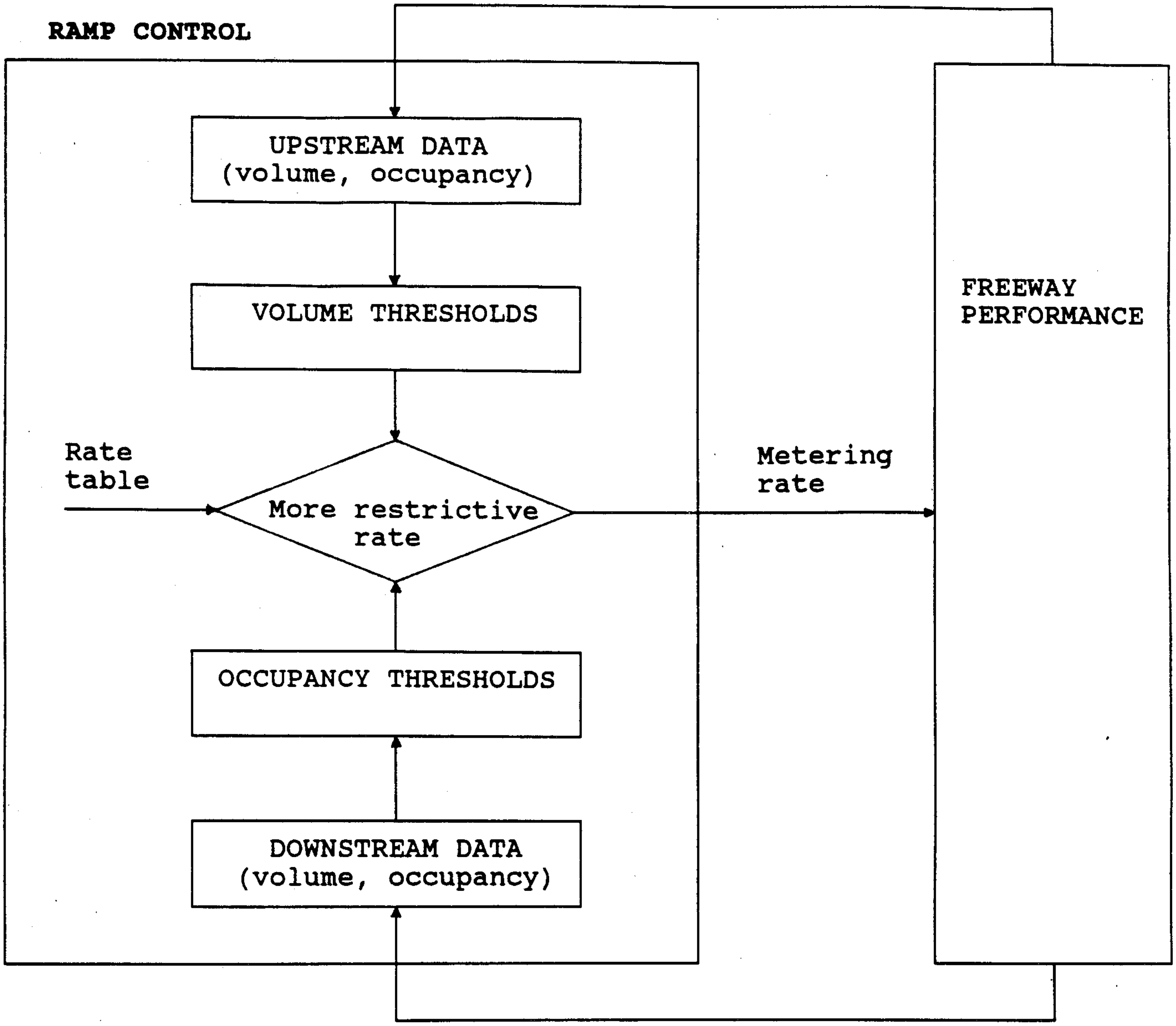

Figure 8 Automatic Rate Selection Process

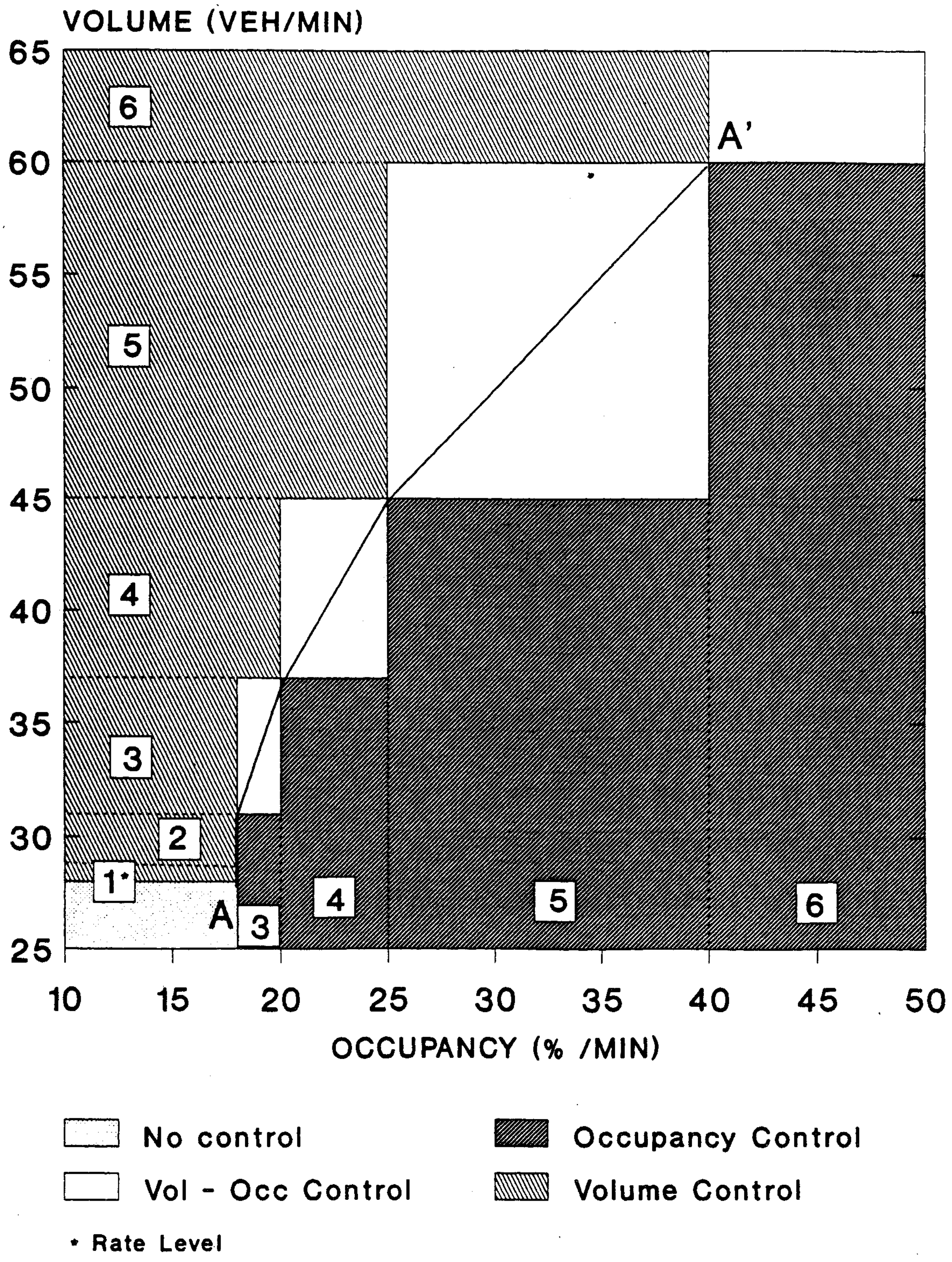

Figure 9. Typical Volume/Occupancy Threshold Policy

IV.3 FREEWAY PERFORMANCE MODULE

Determining the freeway traffic performance with a rigorous and practical model through time and space is a key element in evaluating traffic responsive ramp metering strategies. The model should be able to assess the effects of metering-rate changes on the performance of the freeway mainline and ramps. It should be able to perform such assessment when the changes in metering rates are small. Further, the model should be able to determine the traffic performance at the exact locations of the detector stations. In addition, it should be able to treat changes that occur within a short time interval, such as 6 to 30 seconds, so that it can be transferable across urban traffic operations in which such a rate selection time interval is common.

Considering the above requirements, a macroscopic modeling approach based on simple continuum modeling is adopted. The method discretizes the entire freeway section into short increments, e.g., 100-ft long, as illustrated in Figure 10, and determines the value of traffic parameters in each 100-ft segment every 1 second. The rest of this section describes the modelling methodology adopted in this research.

IV.3.1 Continuum Modelling Methodology

Continuum modelling estimates the density of freeway segments following the vehicle conservation principle and the equilibrium speed-density relationship. The conservation principle can be expressed as follows:

$$K_j(t+\Delta t) = K_j(t) + \frac{\Delta t}{L_j} * \left\{ Q_{j-1,j}(t) - Q_{j,j+1}(t) \right\} \tag{4.0}$$

where $K_j(t)$: density of segment j at time t,

$Q_{j,j+1}(t)$: flow rate from segment j to segment j+1 during time period t, from time t to time t+Δt,

L_j : length of segment j,

Δt : time increment.

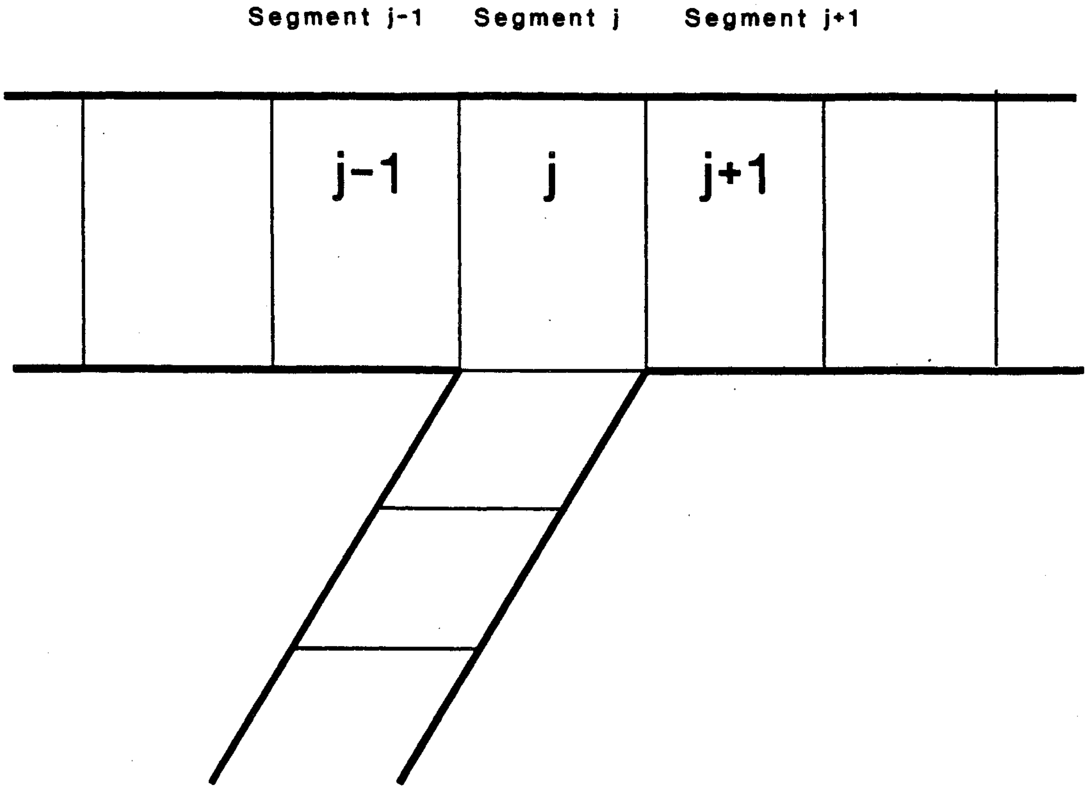

Figure 10 Space Discretization of a Typical Freeway Section

To address the stability issues arising in numerically modeling equation (4.0), we employ an adaptation of the Lax method. A generation term is added to (4.0) to treat the traffic flow at entrance and exit ramps of a freeway. The modified equation that is in agreement with earlier work (Michalopoulos, 1984; Georgadakos et al., 1988) is:

$$K_j(t+\Delta t) = \frac{1}{2}*\left\{K_{j-1}(t)+K_{j+1}(t)\right\} + \frac{\Delta t}{2\Delta X_j}*\left\{Q_{j-1}(t)-Q_{j+1}(t)\right\} + a_j*G_j(t) \tag{4.1}$$

where $K_j(t)$: density of segment j at time t,

$Q_j(t)$: equilibrium flow rate of segment j during time period t (from t to $t+\Delta t$),

$G_j(t)$: generation term of segment j during time period t,

ΔX_j : length of segment j,

Δt : time increment,

a_j : coefficient to convert flow rate into density of segment j.

Our methodology uses the dynamic formulation (4.1) to evaluate the density of each freeway segment at the beginning of each time slice. For obtaining reasonably accurate results, the length of each segment is set at 100 feet. To satisfy stability conditions, the time slice is set at one second.

In Equation 4.1, the equilibrium flow rate Q (vehicles per hour) is defined according to the basic relationship of the flow rate, speed and density:

$$Q_j(t) = K_j(t) * U_j(t) \tag{4.2}$$

where $U_j(t)$ is the space mean speed (miles per hour) in segment j during time period t. This speed is estimated by interpolating the equilibrium speed-density curve:

$$U_j(t) = U_e(\ K_j(t)\ ) \tag{4.3}$$

where $U_e(K_j(t))$ is a relationship that can be used to obtain the speed $U_j(t)$ corresponding to density $K_j(t)$ at equilibrium. During implementation, users of this methodology must specify speed-density relationships calibrated with measured data.

From Equations 4.2 and 4.3, the equilibrium speed and flow rate of each road segment during the present time interval can be calculated when density at the beginning of the time interval is known. The density at the beginning of the next time interval can, then, be calculated from Equation 4.1. Through this iterative calculation process, the traffic parameters, density, speed, and flow rate, of each road segment can be obtained for each time interval.

The vehicle density in a standard road segment that connects with only one upstream and one downstream segments, such as segments j-1 and j+1 in Figure 10, can be calculated by simply setting the generation term in Equation 4.1 to zero. However, special treatment is necessary for modeling system components that connect the freeway to arterials in a network, such as segment j in Figure 10. This treatment is formulated by linking the connecting components, such as a freeway and the connected ramp segment, with the generation term in Equation 4.1. The rest of this section describes the methodologies developed to treat metering, diverging, and entrance-ramp areas.

IV.3.2 Merging at Entrance Ramps

In a freeway merging area (see Figure 11), traffic interaction between two connected segments, a freeway and a ramp, are expressed via two continuum equations, Equation 4.4 and 4.5, with a generation term as a linkage. Equation 4.4 models the density changes on the freeway segment, and Equation 4.5 models the density changes on the corresponding ramp segment:

$$K_j(t+\Delta t) = \frac{1}{2}*\left\{K_{j-1}(t)+K_{j+1}(t)\right\} + \frac{\Delta t}{2\Delta X}*\left\{Q_{j-1}(t)-Q_{j+1}(t)\right\} + a_j*Q_{in}(t) \tag{4.4}$$

$$K_{j,r}(t+\Delta t) = \frac{1}{2}*\left\{K_{j-1,r}(t)+K_{j,r}(t)\right\} + \frac{\Delta t}{2\Delta X}*\left\{Q_{j-1,r}(t)+Q_{j,r}(t)\right\} - a_{j,r}*Q_{out,r}(t) \tag{4.5}$$

where subscript r indicates an entrance-ramp segment.

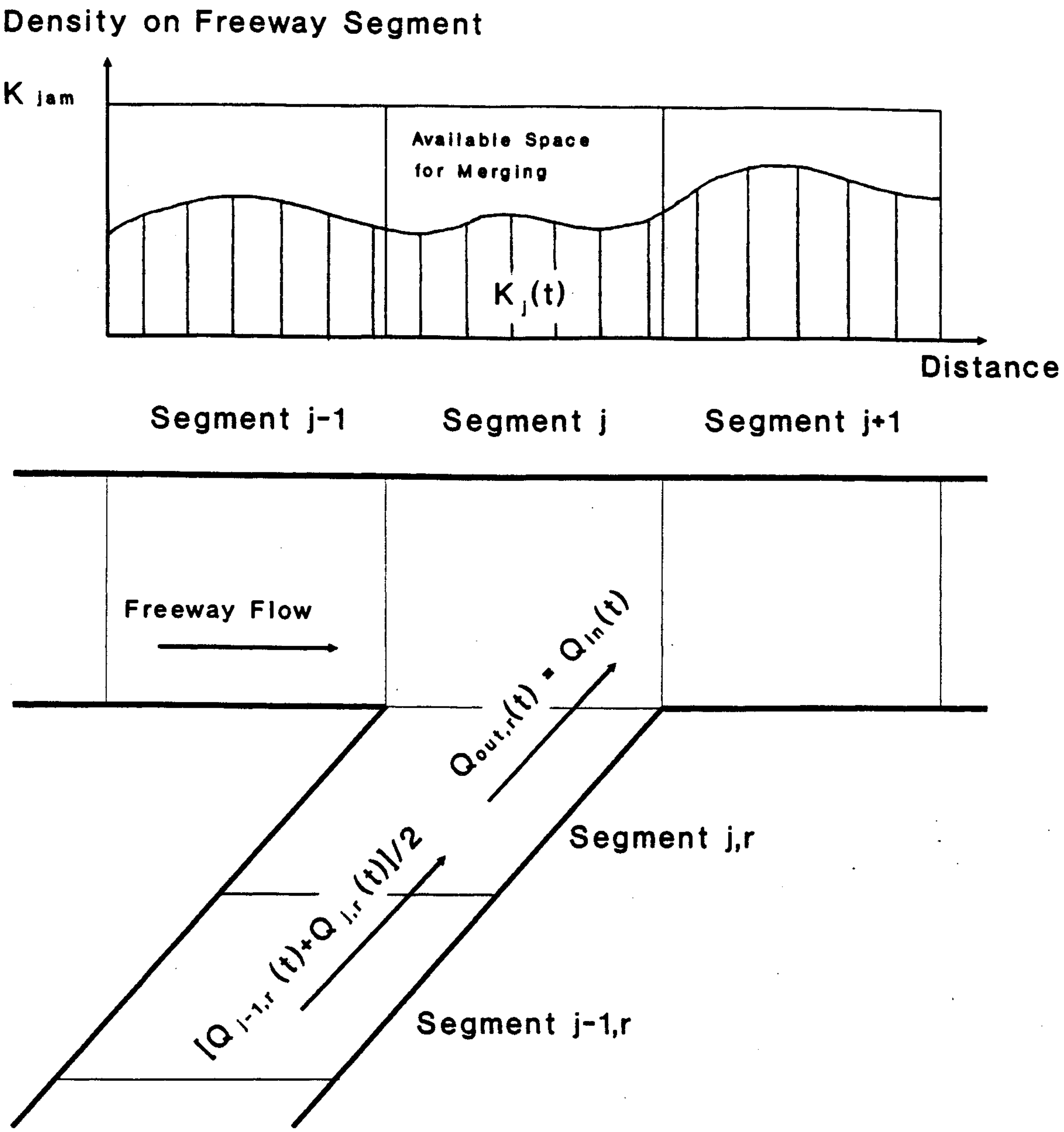

Figure 11 Ramp Merging Flow

The generation term is defined as the actual flow rate merging at an entrance ramp, and is equal to the flow rate $Q_{out,r}(t)$ leaving ramp segment j and the flow rate $Q_{in}(t)$ entering the corresponding freeway segment j that are denoted in Equations 4.4 and 4.5 respectively. The actual merging rate is defined as the minimum value of desired merging rate and limited merging rate. Desired merging flow rate equals the current ramp flow rate $Q_{j,r}(t)$ which can be estimated by Equation 4.2:

$$\text{desired } Q_{out,r}(t) = Q_{j,r}(t) = K_{j,r}(t) * U_{j,r}(t) \tag{4.6}$$

Further, limited merging flow rate is defined as:

$$\text{limited } Q_{out,r}(t) = \{ K_{j,jam} - K_j(t) \} * U_j(t) \tag{4.7}$$

where $K_{j,jam}$ is the jam density of freeway section j. The difference between jam density and the current density is the available storage space on a freeway segment for merging. The product of the available density and the current speed expresses the highest flow rate allowed for merging into the freeway segment.

Thus, actual merging flow rate is:

$$\begin{aligned} Q_{in}(t) &= Q_{out,r}(t) \\ &= \min \{ \text{desired } Q_{out,r}(t), \text{ limited } Q_{out,r}(t) \} \end{aligned} \tag{4.8}$$

On ramp segment j,r, the entering flow from upstream is defined as the average of the equilibrium flow of upstream and the ramp segment, i.e.

$(Q_{j-1,r}(t)+Q_{j,r}(t))/2$. This upstream entering flow and the downstream merging flow satisfy the conservation principle. In addition, the downstream density is assumed equal to the density of the ramp segment, i.e. $K_{j+1,r}(t) = K_{j,r}(t)$. This assumption guarantees stability of the solution.

IV.3.3 Diverging at Exit Ramps

Diverging flow at exit ramps (see Figure 12) is modeled as in the case of merging flow. To be sure, actual diverging flow is a fraction of the freeway flow and is limited by the available space on the exit ramp. Therefore, the diverging flow equations are:

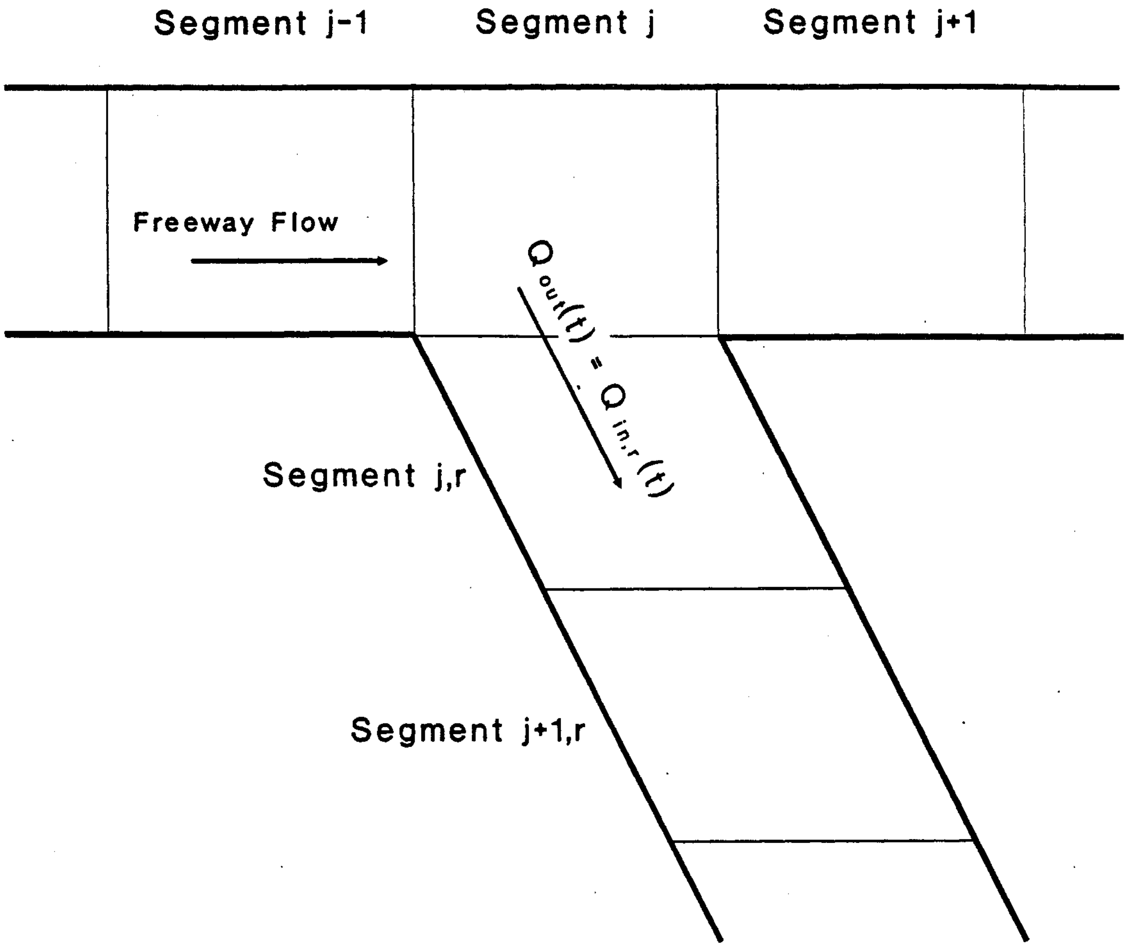

Figure 12. Diverging Flow

$$K_j(t+1) = \frac{1}{2} * \left\{ K_{j-1}(t) + K_{j+1}(t) \right\} + \frac{\Delta t}{2\Delta X} * \left\{ Q_{j-1}(t) - Q_{j+1}(t) \right\} - a_j * Q_{out}(t) \tag{4.9}$$

$$K_{j,r}(t+1) = \frac{1}{2} * \left\{ K_{j,r}(t) + K_{j+1,r}(t) \right\} + \frac{\Delta t}{2\Delta X} * \left\{ Q_{j,r}(t) + Q_{j+1,r}(t) \right\} + a_{j,r} * Q_{in,r}(t) \tag{4.10}$$

The desired diverging flow rate becomes

$$\text{desired } Q_{out}(t) = D(t) * Q_j(t) \tag{4.11}$$

where D(t) is the freeway diverging rate during time period t, and $Q_j(t)$ is the flow rate of freeway segment j. Moreover, the limited diverging flow rate is:

$$\text{limited } Q_{out}(t) = \{ K_{j,r,jam} - K_{j,r}(t) \} * U_{j,r}(t) \tag{4.12}$$

The actual diverging flow rate of segment j is the minimum value of desired diverging flow rate and limited diverging flow rate:

$$\begin{aligned} Q_{in,r}(t) &= Q_{out}(t) \\ &= \min \{ \text{desired } Q_{out}(t), \text{ limited } Q_{out}(t) \} \end{aligned} \tag{4.13}$$

On ramp segment j,r, the leaving flow to downstream is defined as the average of the equilibrium flow of downstream and the ramp segment, i.e. $(Q_{j,r}(t)+Q_{j+1,r}(t))/2$, and the upstream density is assumed equal to the density of the ramp segment, i.e. $K_{j-1,r}(t) = K_{j,r}(t)$.

IV.3.4 Flow Crossing the Stop-Line at Controlled Entrance Ramps

The ramp metering rate of each controlled entrance ramp is obtained from the control module, and is used to estimate the actual metering flow. Figure 13 illustrates road segments adjacent to a ramp meter. The ramp-meter stop line separates the ramp area into two sub-systems. The actual flow rate crossing this stop line links the two sub-systems and is designed as a generation term in the equations of upstream segment j and downstream segment j+1. The actual flow rate should not exceed 1) the ramp metering rate if the upstream demand is greater than the metering rate, 2) the desired flow rate from the upstream segment if demand is less than the metering rate, and 3) the available space of the downstream segment if there is a spillback from the freeway. Therefore, the actual flow rate will be the minimum value of these three constraints. The modified equations are as follows.

For the upstream segment j,

$$K_j(t+1) = \frac{1}{2}*\left\{K_{j-1}(t)+K_j(t)\right\} + \frac{\Delta t}{2\Delta X}*\left\{Q_{j-1}(t)+Q_j(t)\right\} - a_j*R_j(t) \tag{4.14}$$

and for the downstream section j+1,

$$K_{j+1}(t+1) = \frac{1}{2}*\left\{K_{j+1}(t)+K_{j+2}(t)\right\} - \frac{\Delta t}{2\Delta X}*\left\{Q_{j+1}(t)+Q_{j+2}(t)\right\} + a_{j+1}*R_j(t) \tag{4.15}$$

where $R_j(t)$ is the actual flow rate crossing the stop line during time period t. The mathematical expression for $R_j(t)$ is

$$R_j(t) = \min\{\, R_{set}(t),\ Q_{desired}(t),\ Q_{limited}(t)\,\} \tag{4.16}$$

where $R_{set}(t)$: ramp metering rate during time period t.

$Q_{desired}(t)$: upstream desired flow rate during time period t,

$Q_{limited}(t)$: downstream allowed flow rate during time period t.

The desired and limited flow rates are defined as

$$Q_{desired}(t) = \begin{cases} K_j(t) * U_j(t), & \text{if } K_j(t) < K_{j,Qmax} \\ Q_{j,max}, & \text{if } K_j(t) \geq K_{j,Qmax} \end{cases} \tag{4.17}$$

$$Q_{limited}(t) = \begin{cases} Q_{j+1,max}, & \text{if } K_{j+1}(t) \leq K_{j+1,Qmax} \\ K_{j+1}(t)*U_{j+1}(t), & \text{if } K_{j+1}(t) > K_{j+1,Qmax} \end{cases} \tag{4.18}$$

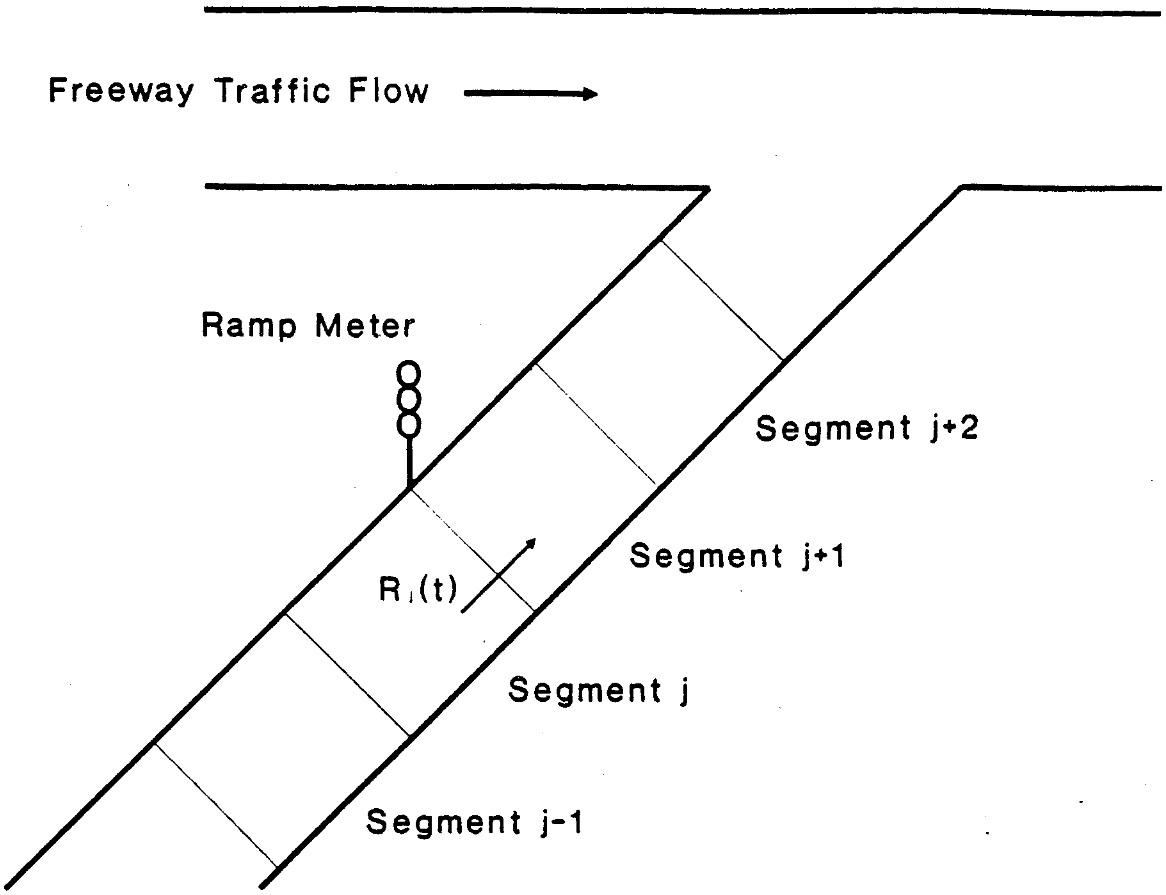

Figure 13. Metering Flow at Entrance Ramp

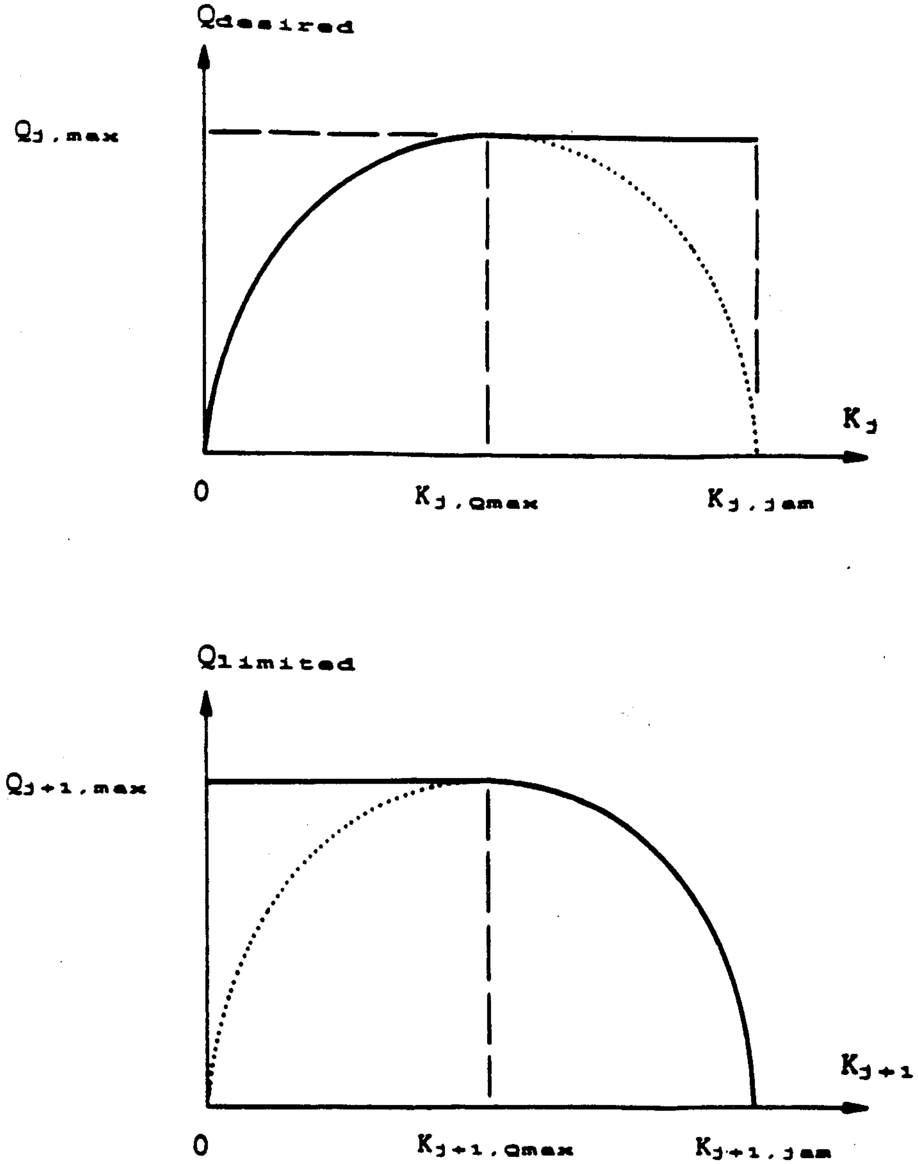

Figure 14. Flow-Density Relationship for Determining Desired and Limited Flow Rate in a Ramp Control Area

where $Q_{j,max}$ is the maximum flow rate of segment j, and $K_{j,Qmax}$ the density corresponding to that flow or "maximum-flow" density. According to these definitions, the flow-density relationships for determining desired and limited flow rates on a ramp are shown in Figure 14. These relationships consider both congestion and non-congestion traffic conditions. For instance, during time period t, when the density of upstream segment j is greater than the maximum-flow density, a queue occurs and the vehicles in the queue desire to be discharged at the maximum flow rate. At downstream segment j+1, when density is less than maximum-flow density, the available space is large enough to allow all vehicles entering the segment at the maximum flow rate. In other traffic situations, the desired and limited flow rates can be obtained from the original Equation 4.2.

IV.3.5 Upstream Demand

At the upstream of the main freeway and each entrance ramp, a dummy segment is added to store vehicles waiting for entering the specified freeway system (see Figure 15). The vehicles are accumulated in this dummy segment and released into the system based on the traffic conditions at the system boundary.

When the densities in both the dummy segment and the upstream boundary segment j are lower than the maximum flow density, traffic is not congested. Assuming that the traffic conditions at the upstream of the dummy segment are the same as the conditions at the dummy segment, the density of the dummy segment is calculated from:

$$K_{dm}(t+\Delta t) = \frac{1}{2} * \left\{ K_{dm-1}(t) + K_j(t) \right\} + \frac{\Delta t}{2\Delta X} * \left\{ Q_{dm-1}(t) - Q_j(t) \right\} \tag{4.19}$$

$$K_{dm-1}(t) = K_{dm}(t), \quad \text{and} \quad Q_{dm-1}(t) = Q_{dem}(t) \tag{4.20}$$

where $K_{dm-1}(t)$ and $Q_{dm-1}(t)$ are the density and flow rate, respectively, at the upstream of the dummy segment, and $K_{dm}(t)$ is the density of the dummy segment. $Q_{dem}(t)$ is the demand flow rate, assumed to be known.

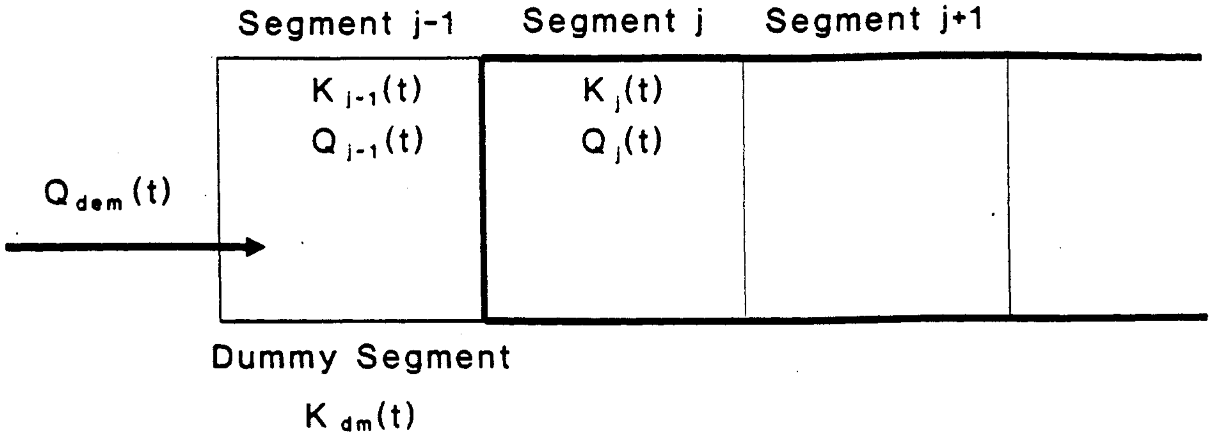

Figure 15. Demand at Upstream Boundary

The density of the upstream boundary segment $K_j(t+\Delta t)$ is calculated from Equation 4.21. In the non-congested case, the upstream density and flow rate of segment j is assumed equal to the density of the dummy segment:

$$K_j(t+\Delta t) = \frac{1}{2} * \left\{ K_{j-1}(t) + K_{j+1}(t) \right\} + \frac{\Delta t}{2\Delta X} * \left\{ Q_{j-1}(t) - Q_{j+1}(t) \right\} \tag{4.21}$$

$$K_{j-1}(t) = K_{dm}(t), \quad \text{and} \quad Q_{j-1}(t) = Q_{dm}(t) \tag{4.22}$$

If traffic is congested, the density of the dummy segment is found from:

$$K_{dm}(t+\Delta t) = K_{dm}(t) + \frac{\Delta t}{\Delta X} * \left\{ Q_{dem}(t) - Q_j(t) \right\} \tag{4.23}$$

Since the density average of upstream and downstream segments is not used in this formulation, vehicles are directly accumulated in the dummy segment. The density of the upstream boundary segment $K_j(t+\Delta t)$ is obtained from Equation 4.21 using the following logic:

```
if ( ($K_{dm}(t) > K_j$) or ($K_{dm}(t) > K_{jam}$) )
then $K_{j-1}(t) = K_j$,
else $K_{j-1}(t) = K_{dm}(t)$                                    (4.24)
```

If the above condition is satisfied, i.e. the density of the dummy segment is greater than the density of the boundary segment j or the jam density, a queue occurs in the dummy segment. In this case, the upstream density of segment j is set as the density of segment j such that vehicles can be vertically stacked in the dummy segment no matter whether the density is greater than jam density. Otherwise, the upstream density of segment j is set as the density of the dummy segment.

Equations 4.19 to 4.24 are formulated for treating upstream demand independently from the queue that may be propagating from downstream.

IV.3.6 Performance Indices

From density, speed and volume, various performance indices including total travel time, delay and volume entering the freeway, can be determined. Performance indices developed in this research include:

Total travel time (veh-hr)

$$= \sum_{t} \{ \sum_{i} [(CD(i)+ND(i))/2 * X(i) * W(i) / 3600] \}$$

Total travel (veh-miles)

$$= \sum_{t} \{ \sum_{i} [CV(i) * X(i) * W(i) / 3600] \}$$

Total volume entering freeway from ramps (vehicles)

$$= \sum_{t} \{ \sum_{i} [CG(i) / A(i) / 3600] \}$$

Total delay (veh-hr)

$$= \sum_{t} \{ \sum_{i^*} [CD(i^*) * X(i^*) * W(i^*)] / 3600 \}$$

i^* : all freeway segments with traffic speed less than a user-specified critical value.

where, t = simulation time unit (1 second),

$CD(i)$ = current density of segment i (veh/mile),

$ND(i)$ = next-time density of segment i (veh/mile),

$X(i)$ = length of segment i (mile),

$W(i)$ = width of segment i (lane),

$CV(i)$ = current volume of segment i (veh/hr),

$CG(i)$ = volume entering freeway from i th ramp (veh/hr),

$A(i)$ = coefficient to convert volume to density for i th ramp (1/mile).

IV.4 SUMMARY

A control-emulation method for real-time selection of ramp metering rates is developed in this chapter.. The method includes a ramp control module and a freeway performance module. The Ramp Control module emulates the ramp-metering control logic with the traffic data received from the Freeway Performance module, and sends the determined metering rates to that module. Freeway

Performance uses these rates to simulate traffic dynamics on the freeway and ramps for an assumed demand. After a certain period, which is set by the control logic, Freeway Performance sends the simulation results, consisting of simulated data at pre-defined detector locations, to the Ramp Control module. Ramp Control determines metering rates for the next time period, and the iterative process continues until the end of the period.

The on-line automatic rate-selection control strategy currently implemented in the Twin Cities is adopted in the ramp control module as the control logic that determines metering rates. For each ramp and every thirty seconds, the control strategy requires one upstream volume and five downstream occupancy measurements from pre-defined detectors. Upstream volume and the highest downstream occupancy are compared with pre-determined volume/occupancy thresholds. The volume and occupancy rates corresponding to the matched thresholds are identified from a rate table, and the most restrictive rate is selected. This control strategy is exactly emulated by the ramp control module.

The freeway performance module employs continuum modelling to simulate the dynamics of the traffic density on a freeway system following the vehicle conservation principle and the equilibrium speed-density relationship. In particular, this module simulates the traffic dynamics at specific areas of the system, such as ramp-meter, merging, and diverging areas. The module determines the value of traffic variables, such as volume and occupancy, at pre-defined detector locations, and can control ramp flows according to the metering rates determined by the ramp control module. Several traffic performance indices, such as total travel time, total entering volume, and total delay, are provided by the module for evaluating control strategies.

V. TESTING AND APPLICATION OF CONTROL-EMULATION METHOD

V.1 INTRODUCTION

In this chapter, the features of the new control-emulation method are illustrated through application to a freeway in the Minneapolis-St. Paul Metropolitan Area. More specifically, a 6-mile section of the I-35W freeway in Minneapolis was selected as the sample freeway section. This section, located south of the Minneapolis downtown area, contains 18 entrance/exit ramps and a variety of geometric types such as merging, diverging and weaving. The test section is presented in Figure 3, which also includes the loop detector locations, where volume and occupancy measurements can be obtained every minute. The rest of this chapter summarizes the test results of the new method with the real data collected from the test freeway section and an example evaluation of the alternative volume/occupancy threshold policies for the sample freeway section.

V.2 TESTING OF METHOD

The control emulation method was first tested with the real volume data collected at the test site. Based on the current metering strategy, simulated volume was compared with the real data collected from the loop detectors at each detector location. Volume comparison results are presented in Figures 16 and 17 for 5-min data over 3-hr periods on Wednesday, November 14, and Thursday, December 6, 1990. From Table 6, comparison of our results with real volume data indicates mean percentage error ranging from 5 to 19 percent. Performance could improve with improvements in modeling, an issue currently addressed by the authors.

V.3 EXAMPLE EVALUATION OF THRESHOLD POLICIES

V.3.1 Evaluation of Occupancy-Threshold Changes

As an illustrative example, the control emulation method is applied to determining the effects of occupancy threshold changes on traffic performance in a sample freeway section. As part of this example, four alternative threshold policies were formulated by increasing/decreasing occupancy thresholds by 2 and 4 percentage points above/below their current values respectively. The four

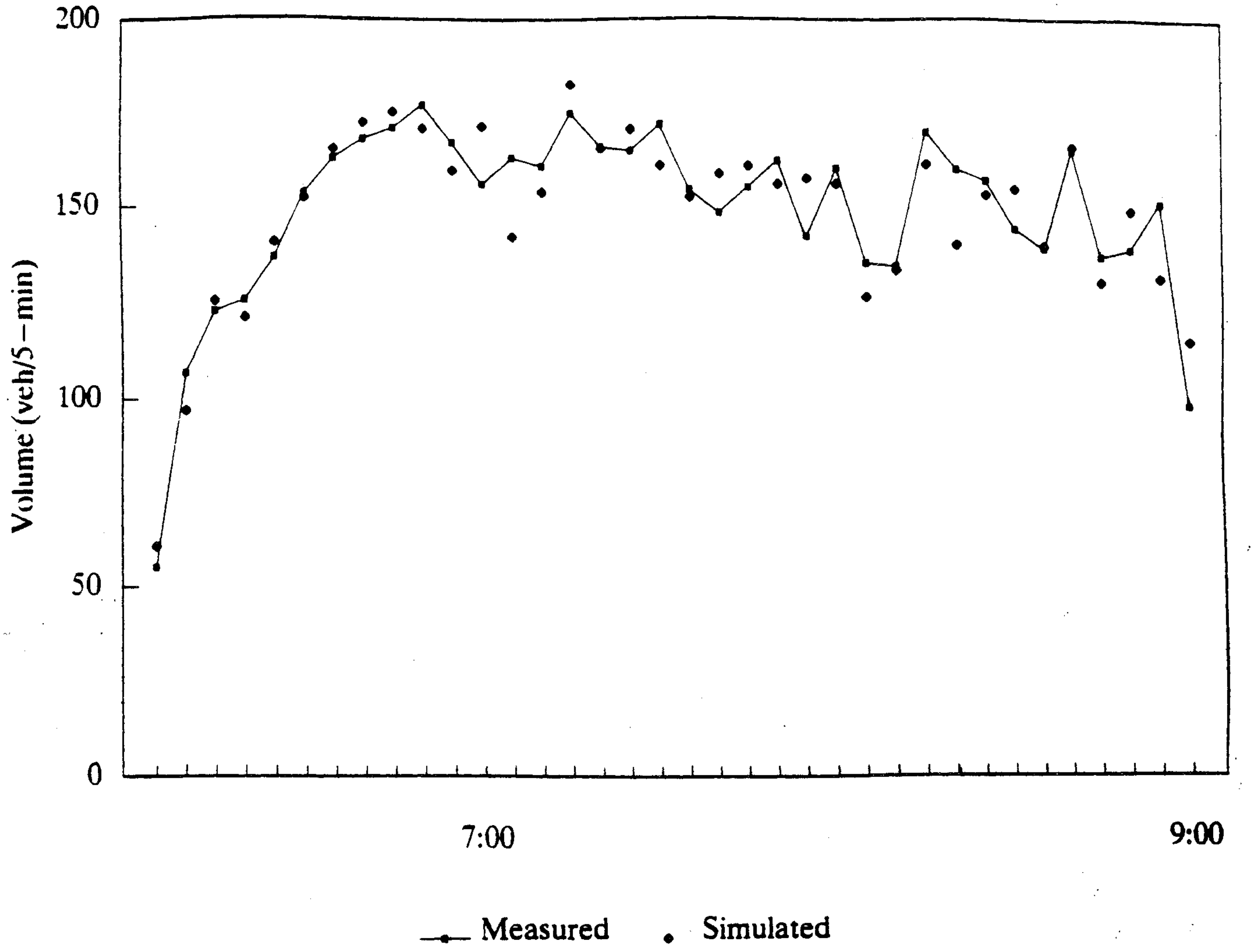

Figure 16. Five-Minute Volume: Station 41, Dec. 6, 1990

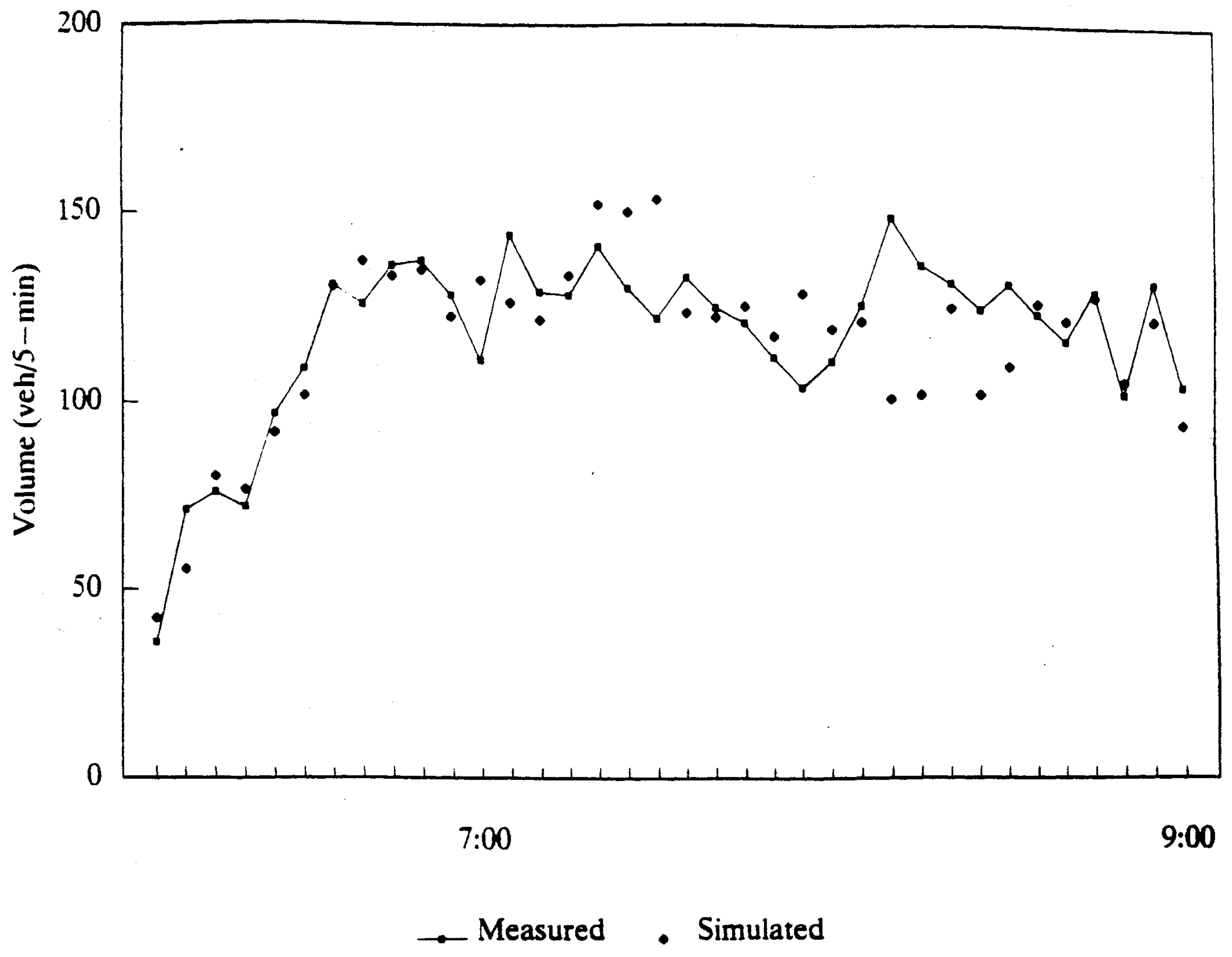

Figure 17. Five-Minute Volume: Station 48, Dec. 6, 1990

Table 6. Mean Percentage Error

Detector Location	Nov. 14	Dec. 6
110N	5 %	5 %
106N	5	5
102N	5	5
98N	6	5
94N	18	13
90N	7	6
86N	7	6
82N	7	8
78N	11	19
76N	9	10
73N	9	9
70N	9	9
66N	9	10
63N	10	11
Total	8 %	9 %

policies and the current one are presented in Figure 18. As the Figure indicates, the volume thresholds of the four policies remain unchanged.

For the purpose of this evaluation and since real entrance ramp demand data were not available, a hypothetical demand pattern, which is higher than normal demand for the sample site, was created as illustrated in Figure 19. Increasing current demand offers a better understanding of the performance of metering policies under varying congestion conditions. Figures 20 - 23 illustrate the effects of occupancy threshold changes on traffic performance at one location (detector station #41) in the test freeway section. The traffic performance points in the congestion area are shifted to the right and up when a lower occupancy-threshold policy is used. For instance, the highest occupancy is reduced from 34% to 31% and the corresponding volume is increased from 30 to 33 vehicles/minute as the occupancy threshold changes from O-22 to O-14. This indicates that the most restrictive threshold policy (O-14) resulted in a smaller number of congested volume-occupancy points, while the least restrictive occupancy threshold (O-22) resulted in a more congested traffic performance.

To evaluate each policy, the following performance indices were selected and their values were assessed.

- Total volume entering the freeway during the simulation period,
- total delay, defined as the total travel time of all vehicles at a speed less than 45 mph,
- total travel times of vehicles on the freeway and in the whole system including vehicles waiting outside the system, and
- average speed for the entire simulation period.

Table 7 and Figures 24-28 summarize the simulation results over a 3-hr period including the no-control case, which represents the "before control" situation. As expected, the no-control case indicates the worst traffic performance in terms of delay, average speed and total travel time on the freeway; however, total volume entering the freeway is the highest. Futher, the results indicate that, at given volume thresholds, traffic performance is more sensitive to occupancy threshold decreases below the current value than to threshold increases. For example, decreasing the current eighteen-percent threshold by 2 percentage points results in a 7.3% increase in average speed, a 17.3% decrease in total delay and a 2.4 % decrease in total volume entering the freeway. However, shifting

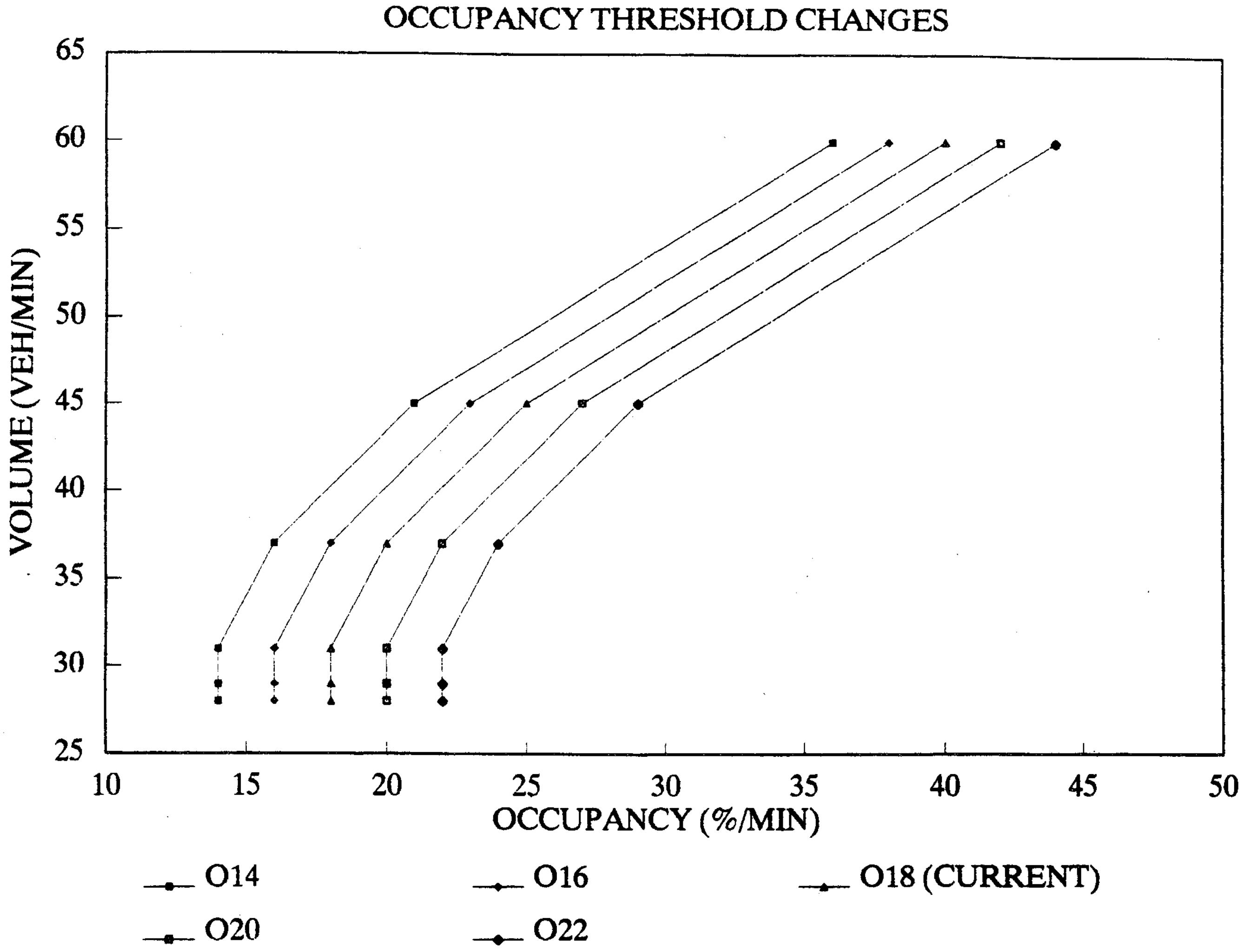

Figure 18. Occupancy-Threshold Strategies

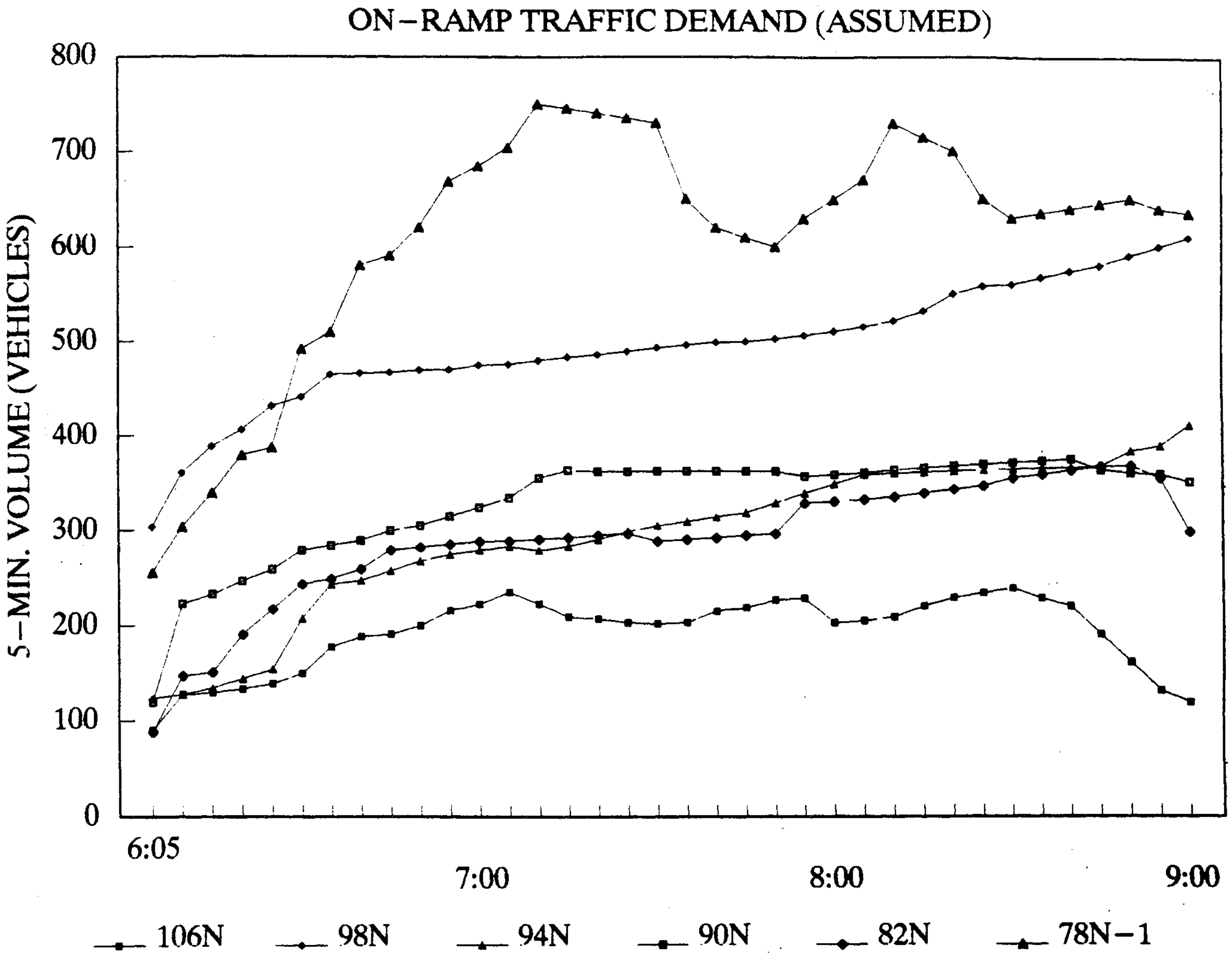

Figure 19. On-Ramp Traffic Demand (Assumed)

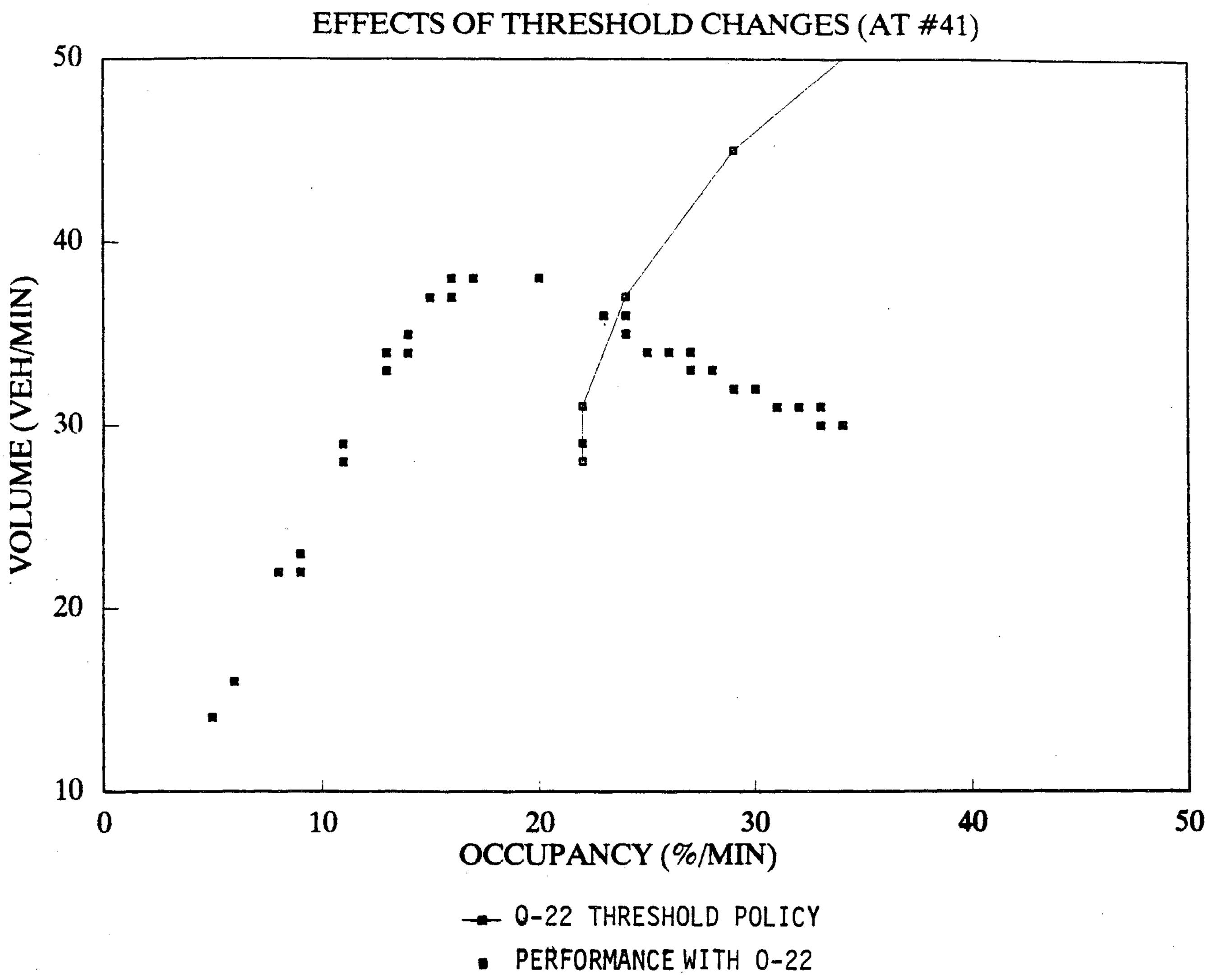

Figure 20. Effects of O-22 Threshold Policy (at #41)

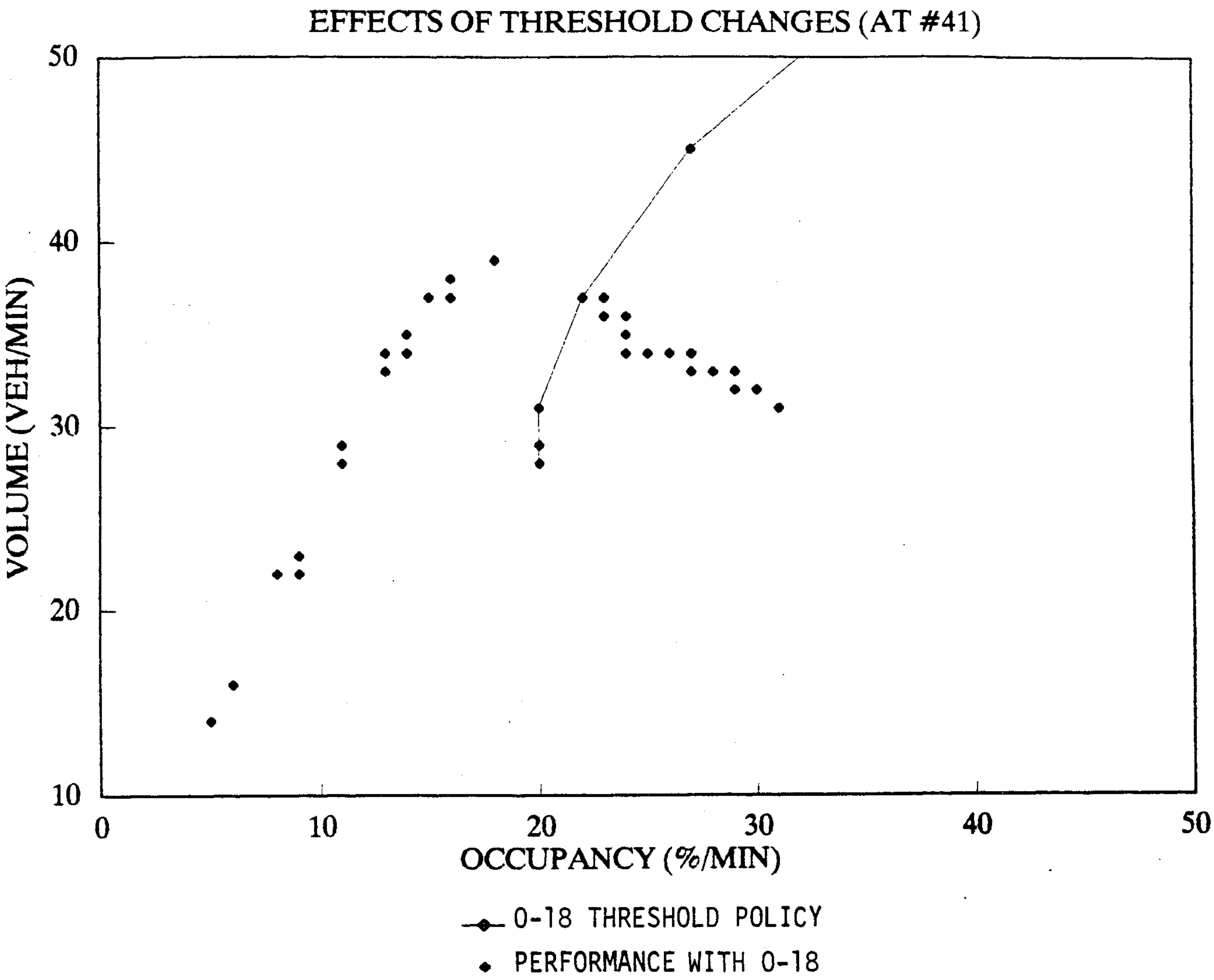

Figure 21. Effects of O-18 Threshold Policy (at #41)

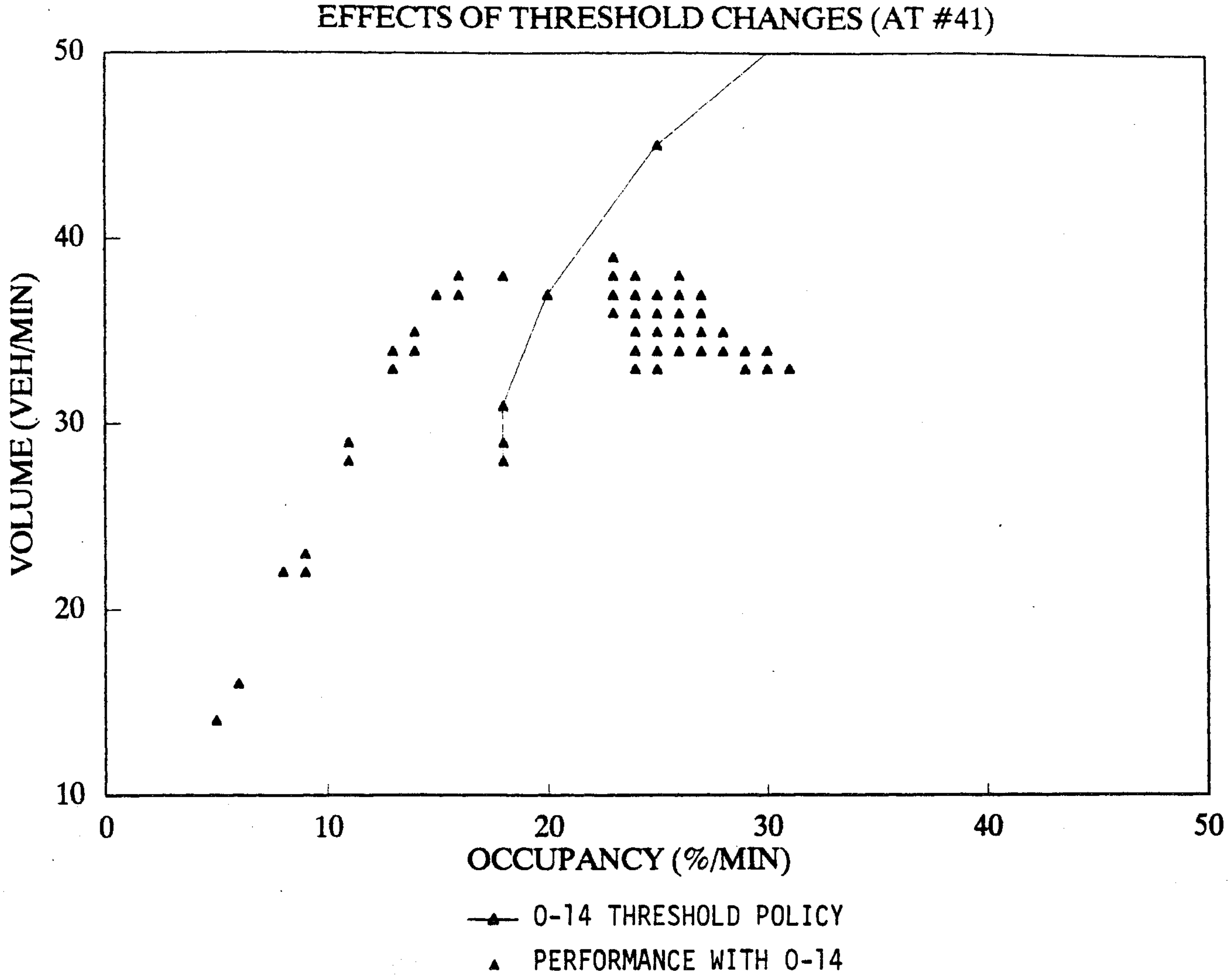

Figure 22. Effects of O-14 Threshold Policy (at #41)

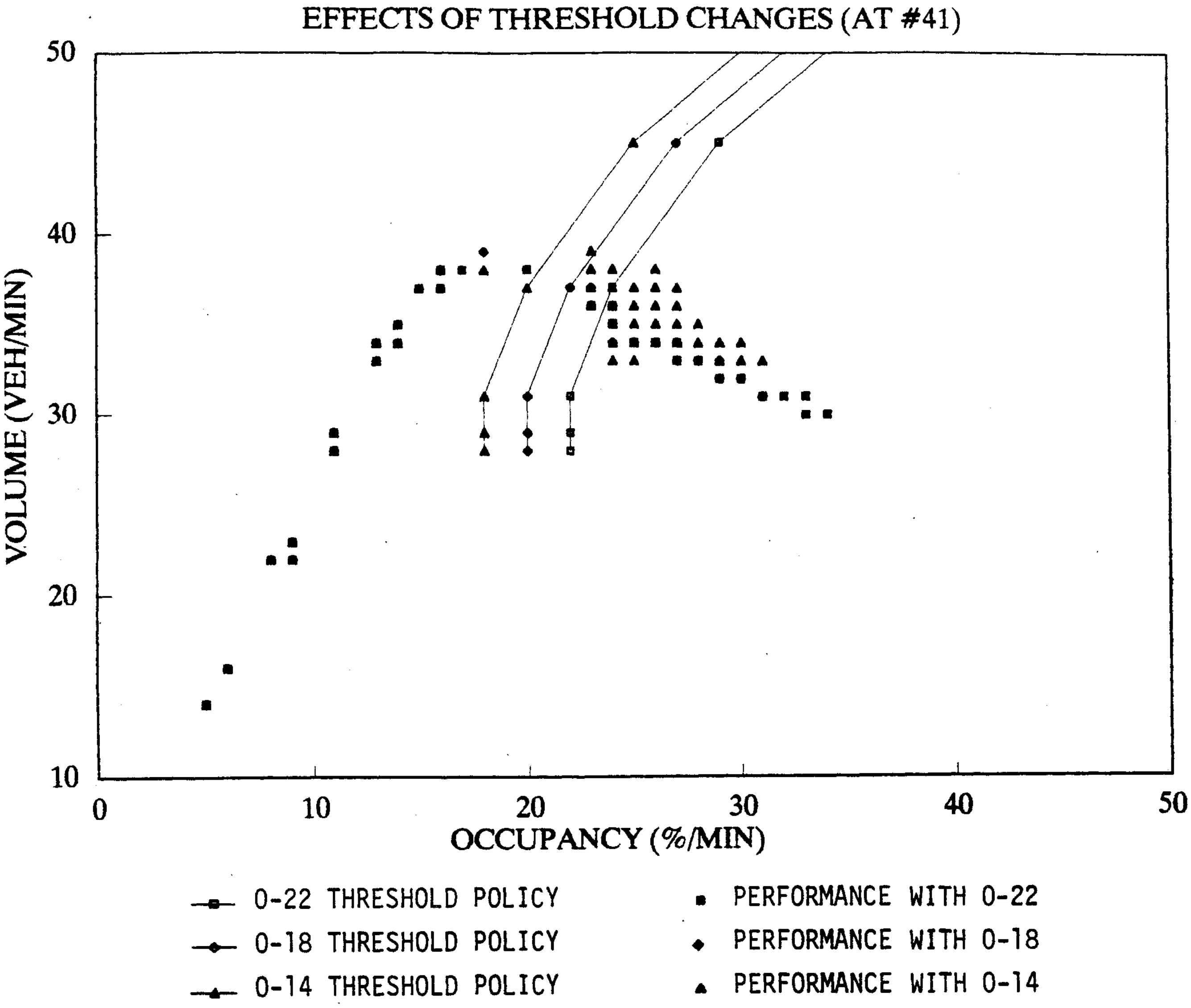

Figure 23. Effects of Occupancy-Threshold Changes (at #41)

Table 7 : Evaluation Results for Occupancy-Threshold Changes

	Total Volume entering from ramps (veh)	Total Delay [speed<45] (veh-hrs)	Average Speed (mph)	Total Travel Time (veh-hrs)		
				Freeway	Ramp	System*
No Control	9320 [21.3]	1428 [37.3]	33 [-21.6]	1849 [19.2]	40 [-81.2]	4555 [1.3]
O-22	7781 [1.3]	1119 [7.6]	39 [-5.4]	1607 [3.6]	177 [-15.7]	4737 [5.4]
O-20	7772 [1.2]	1067 [2.6]	41 [-1.8]	1568 [1.1]	183 [-12.6]	4562 [1.5]
O-18	7683 [0.0]	1040 [0.0]	42 [0.0]	1551 [0.0]	210 [0.0]	4496 [0.0]
O-16	7496 [-2.4]	862 [-17.2]	45 [7.3]	1464 [-5.6]	221 [5.7]	4373 [-2.8]
O-14	7328 [-4.6]	709 [-31.8]	47 [13.1]	1392 [-10.2]	224 [7.0]	4443 [-1.2]
O-12	7384 [-3.9]	404 [-61.1]	52 [25.9]	1264 [-18.5]	249 [18.9]	4311 [-4.1]
O-10	7055 [-8.2]	70 [-93.3]	58 [39.7]	1129 [-27.2]	264 [26.2]	4673 [3.9]

[ ] : Percentage difference with respect to current threshold policy (Occ-18)

* : Including vehicles waiting outside system, i.e. waiting outside the upstream boundary of the freeway and entrance ramps

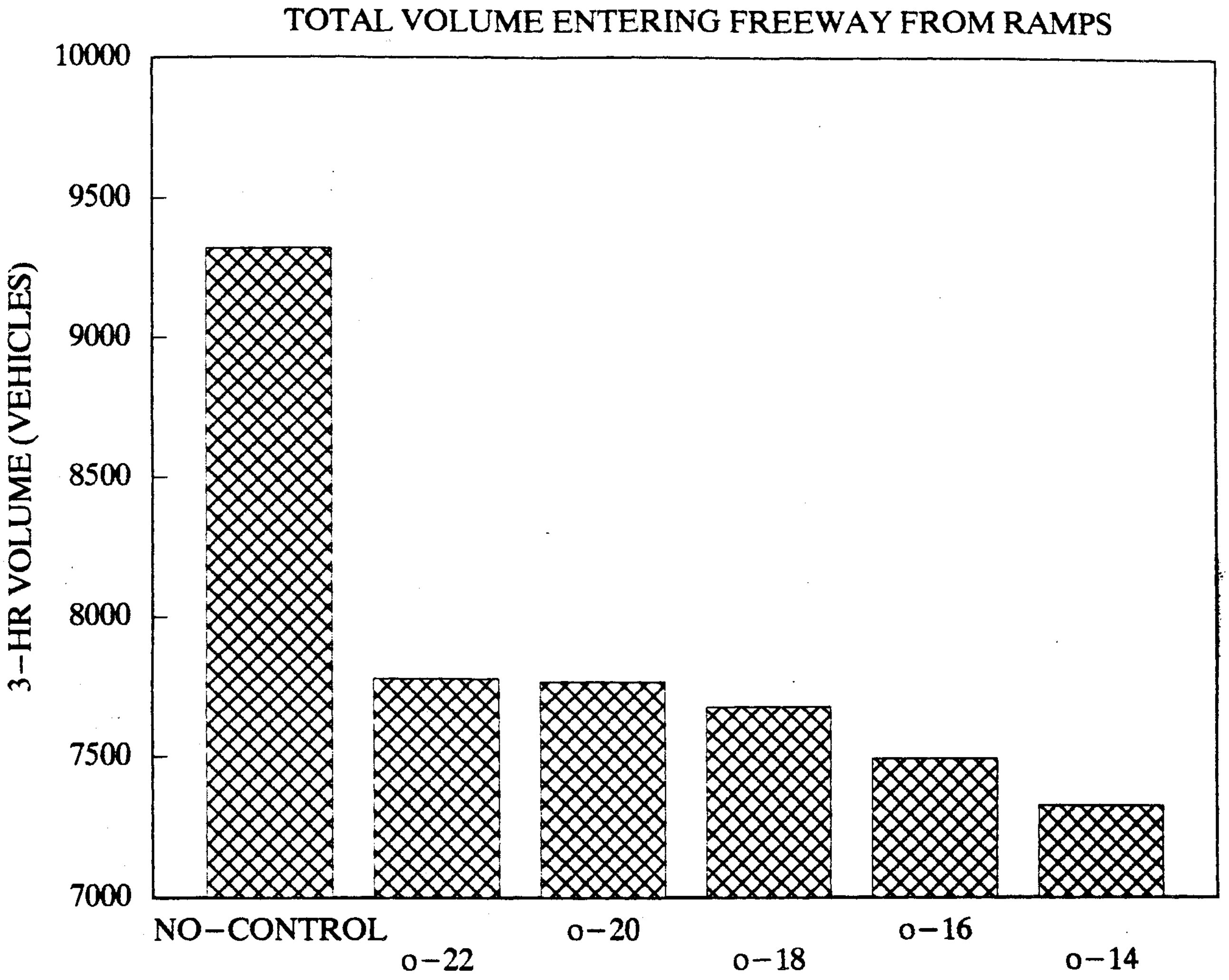

Figure 24. Total Volume Entering Freeway from Ramps for Occupancy-Threshold Changes

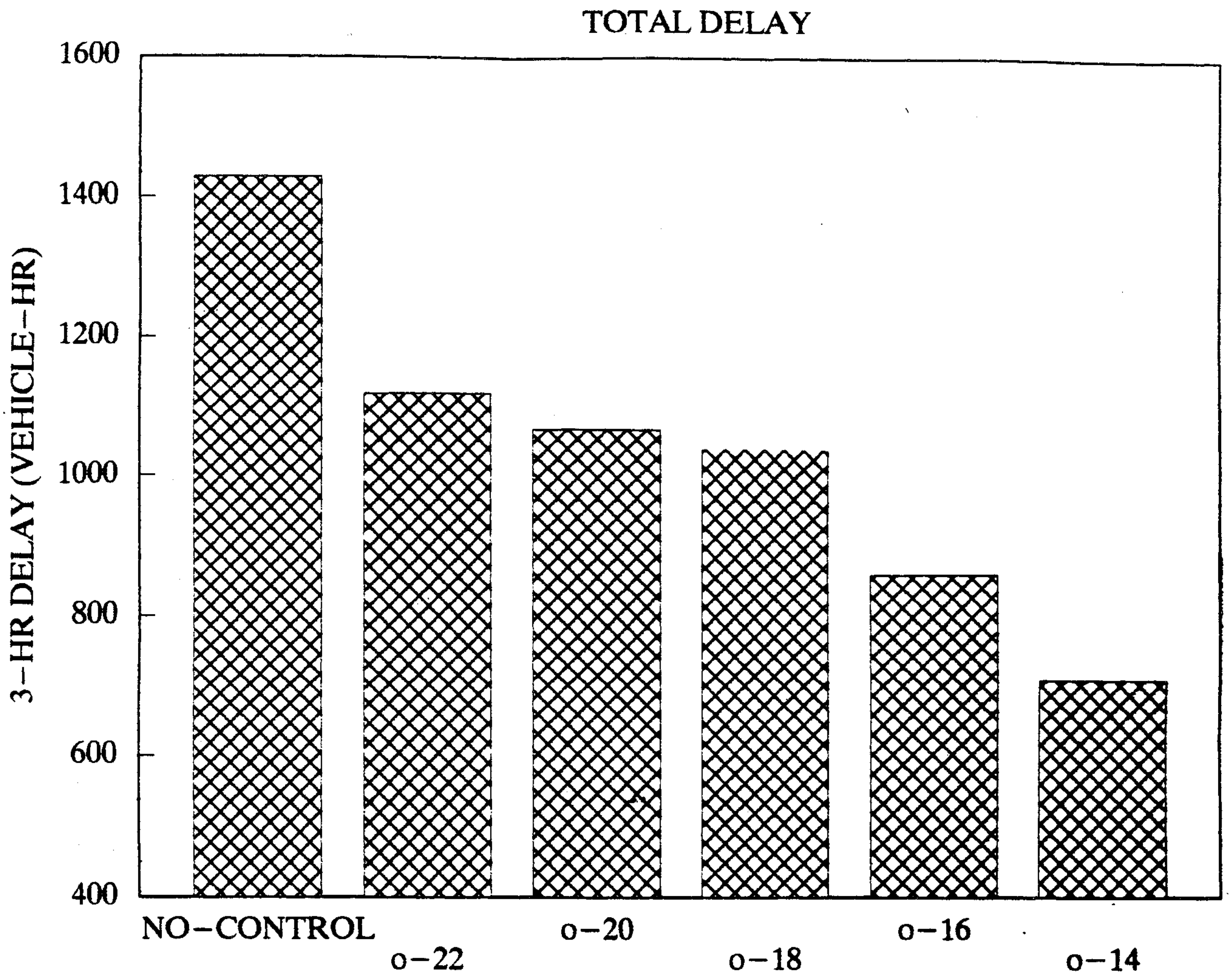

Figure 25. Total Delay for Occupancy-Threshold Changes

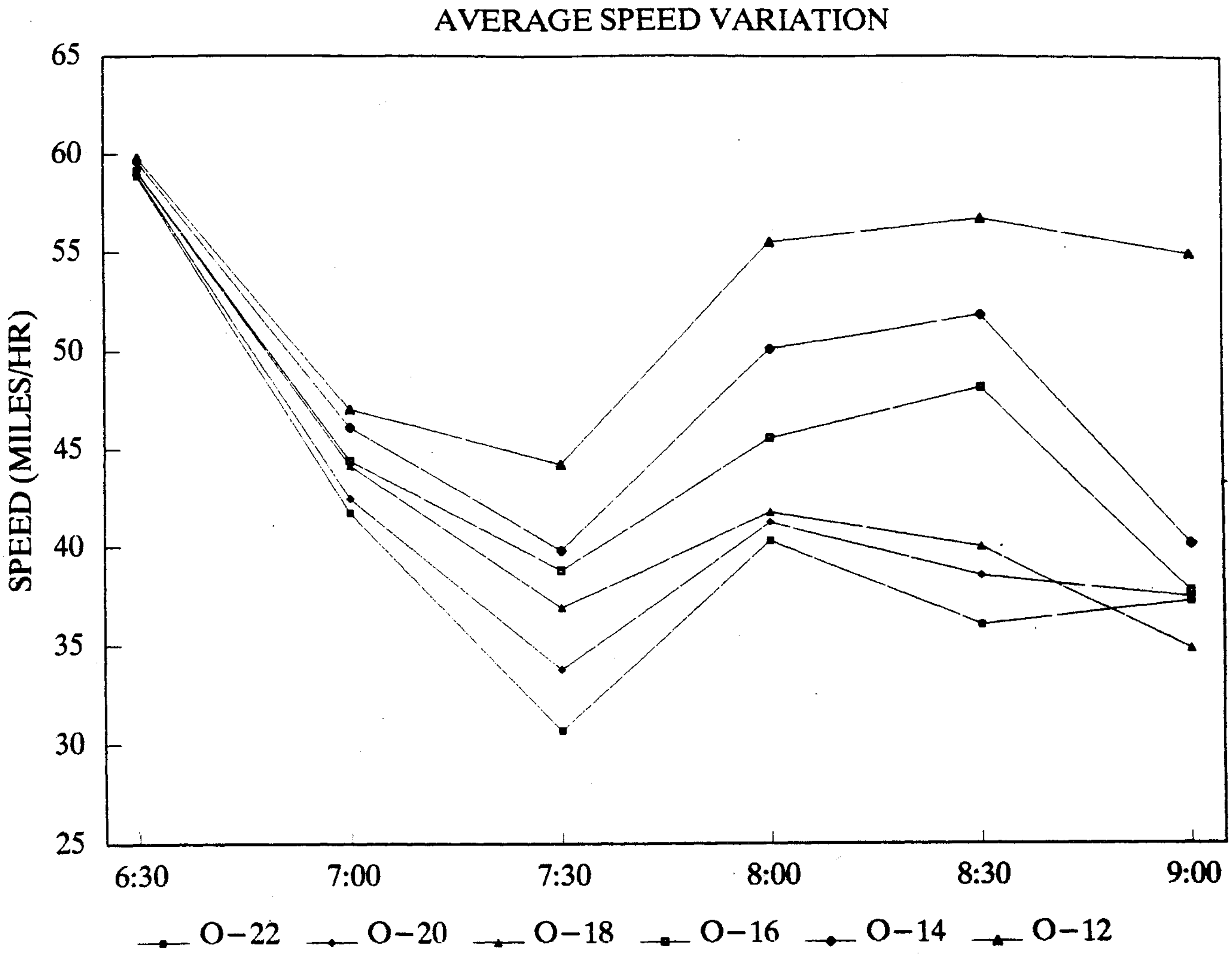

Figure 26. Average Speed Variation for Occupancy-Threshold Changes

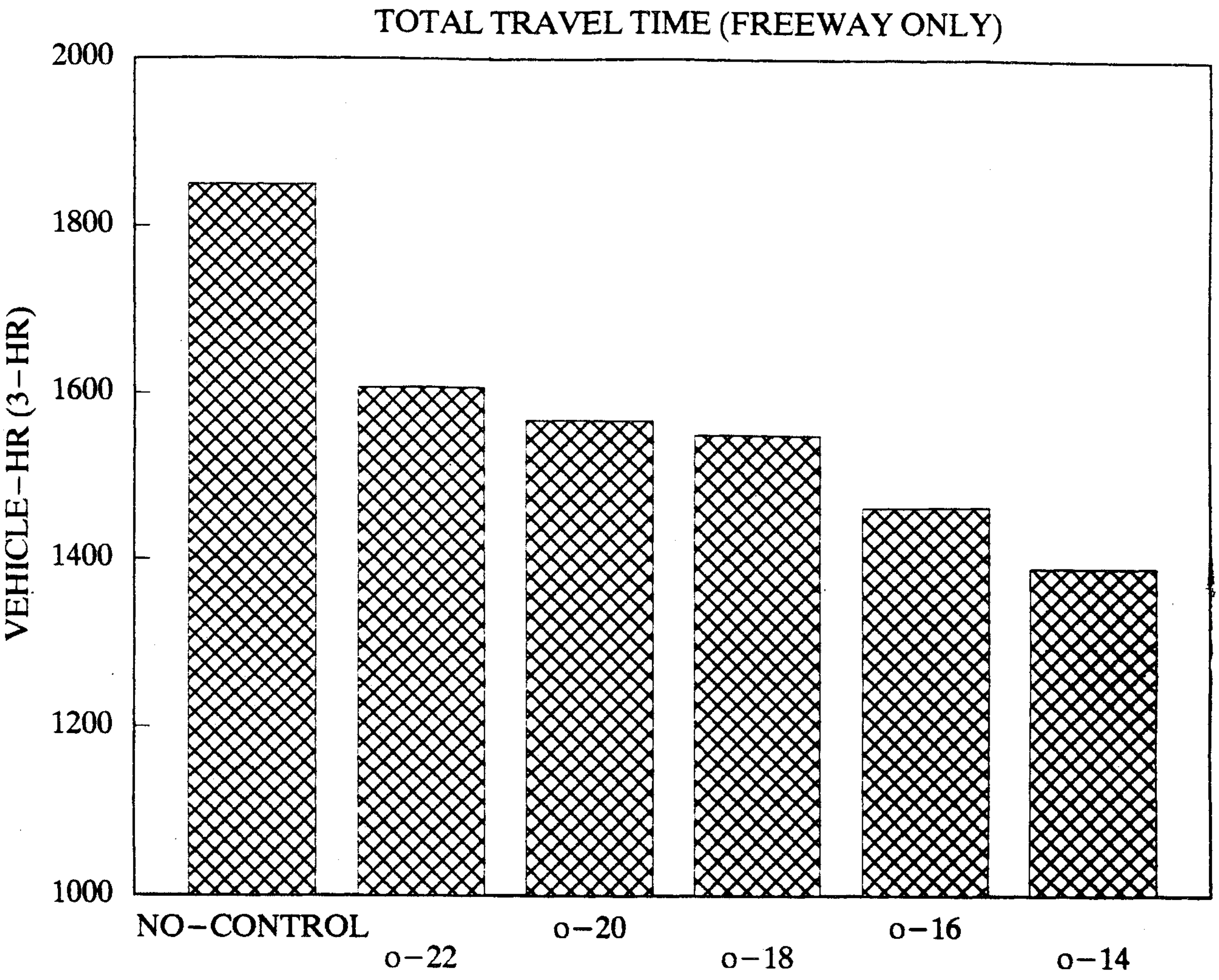

Figure 27. Total Travel Time (Freeway Only) for Occupancy-Threshold Changes

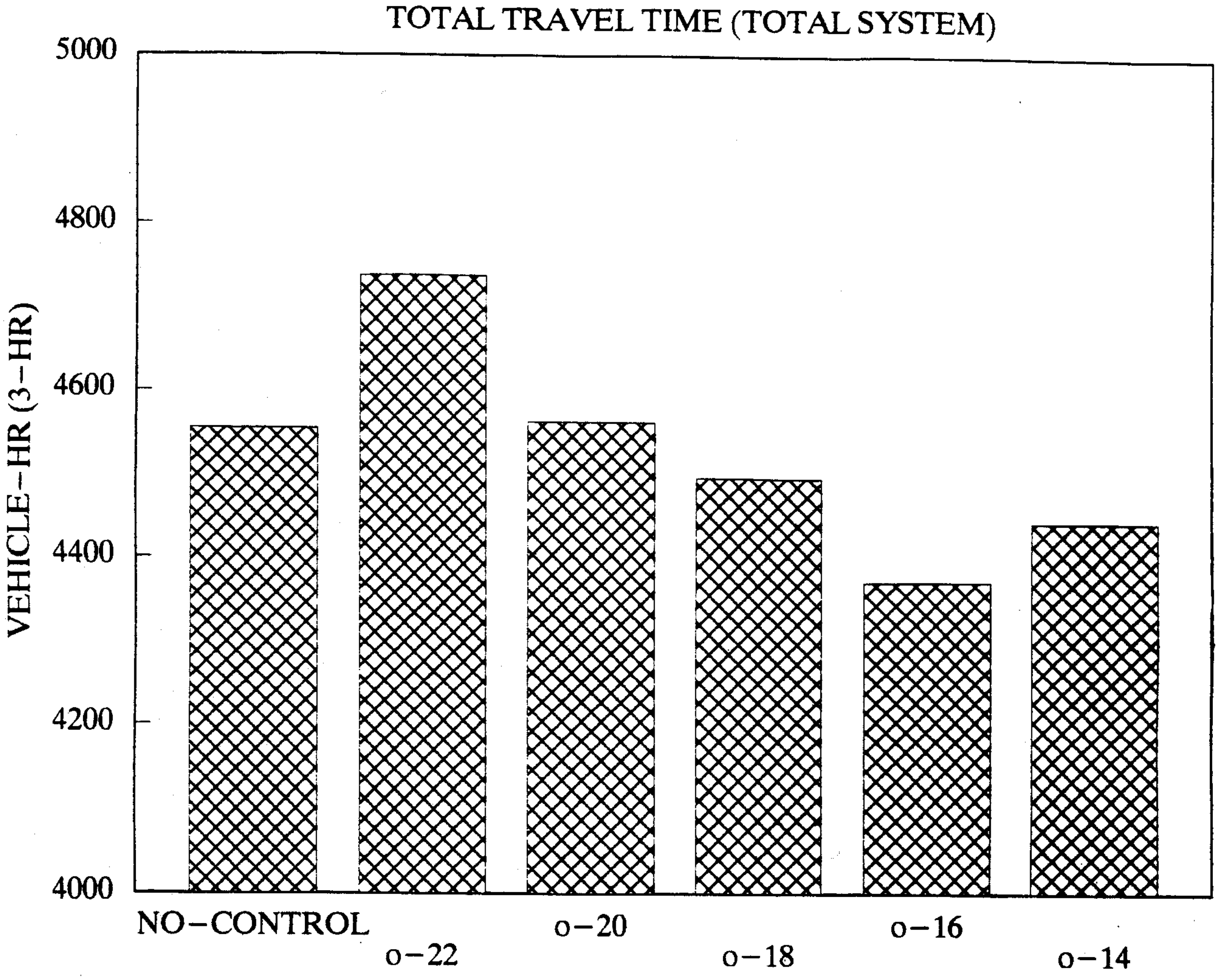

Figure 28. Total Travel Time (Total System) for Occupancy-Threshold Changes

up the current threshold by 2 percentage points results in a 1.8% speed decrease, a 2.5% delay increase and a 1.2% volume increase. This phenomenon is due to the fact that, as occupancy threshold increases, volume thresholds become a more dominant factor in determining metering rates. Since volume thresholds remain constant in this application example, the effect of occupancy thresholds on traffic performance becomes less significant as occupancy threshold values increase.

V.3.2 Evaluation of Volume-Threshold Changes

Using the same demand pattern and performance indices as in the previous section, the effects of volume-threshold changes were evaluated. Figure 29 summarizes the alternative threshold policies formulated by increasing/decreasing the current volume thresholds by 3 and 6% respectively, while the occupancy thresholds remain unchanged. Table 8 and Figures 30-33 present the simulation results over a 3-hr period for the test site. The results suggest that, similar to the occupancy changes, increasing volume-thresholds above their current value has little impact on traffic performance, while decreasing volume thresholds results in significant improvements in terms of delay and average speed with relatively small reduction in the total volume entering the freeway. For example, a 3% decrease in volume thresholds resulted in a 21% speed increase, a 51% delay reduction and a 4.8% volume decrease; however, a 3% increase in volume thresholds does not have a significant impact on traffic performance.

The above simulation results are summarized in the delay-volume space in Figure 34 and 35, which indicates that, for the given freeway section and demand pattern, employing a more restrictive metering policy could substantially reduce the total delay on the freeway while incurring a small relative decrease in the total volume using the freeway. By considering the trade-offs between volume increase and delay reduction, a desirable threshold policy can be determined for a given freeway section.

V.4 SUMMARY

In this chapter, the new control-emulation method was tested in emulating volume and occupancy thresholds, rate tables and automatic rate-selection control strategies on I-35W in Minneapolis. The tests indicated errors in the range from 5 to 12 percent. The method was also applied to assess the sensitivities of volume/occupancy threshold changes on traffic performance with hypothetical demand data. The simulation results were in the expected direction, and indicated the ability to evaluate the effects of small changes in metering policies on traffic performance.

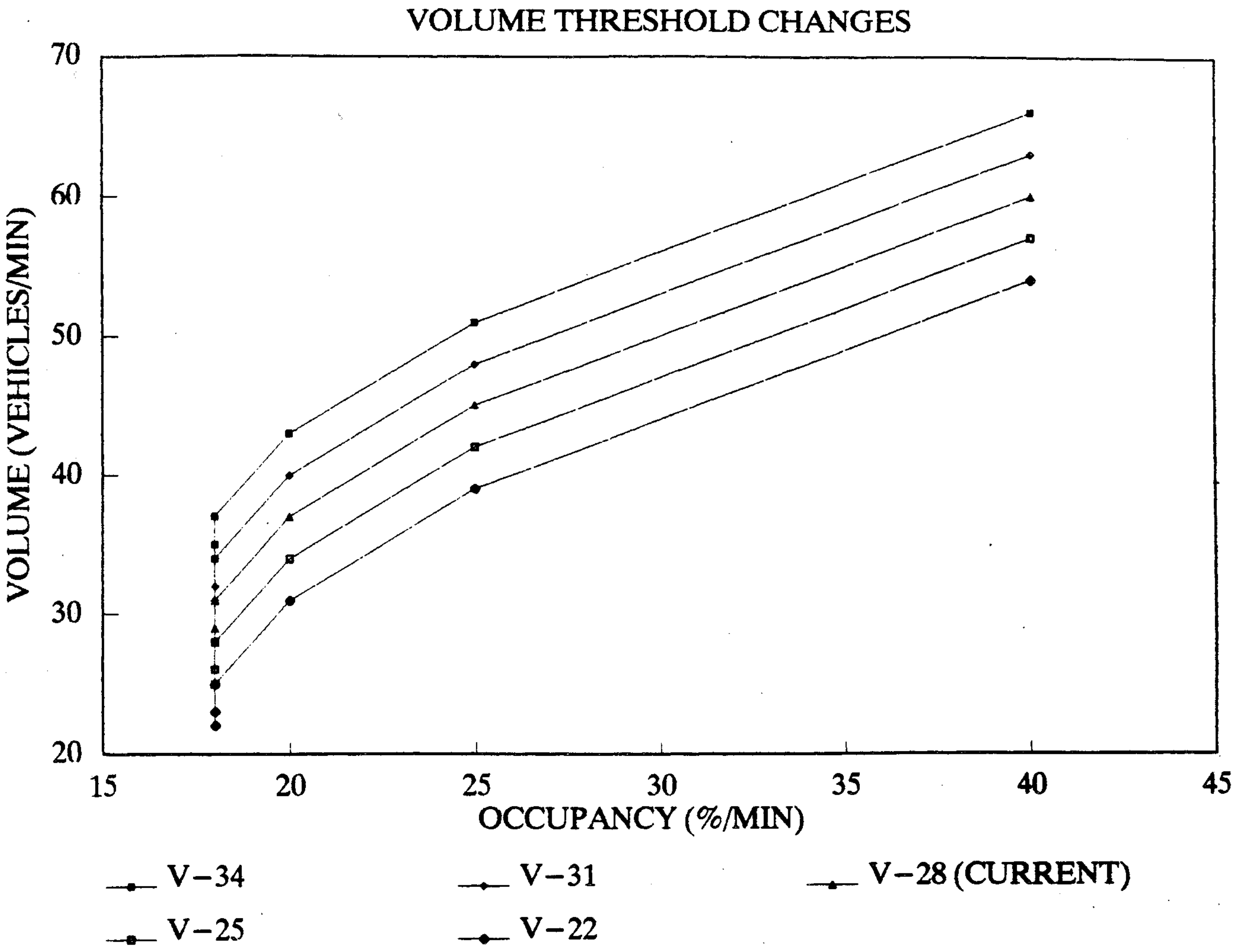

Figure 29. Volume-Threshold Strategies

Table 8: Evaluation Results for Volume-Threshold Changes

	Total Volume entering from ramps (veh)	Total Delay [speed<45] (veh-hrs)	Average Speed (mph)	Total Travel Time (veh-hrs)		
				Freeway	Ramp	System*
No Control	9320 [21.3]	1428 [37.3]	33 [-21.6]	1849 [19.2]	40 [-81.2]	4555 [1.3]
V-34	7814 [1.7]	1049 [0.9]	41 [-0.1]	1558 [0.5]	199 [-5.2]	4450 [-1.0]
V-31	7826 [1.9]	1049 [0.9]	41 [-0.1]	1559 [0.5]	203 [-3.2]	4456 [-0.9]
V-28	7682 [0.0]	1040 [0.0]	42 [0.0]	1551 [0.0]	210 [0.0]	4496 [0.0]
V-25	7315 [-4.8]	506 [-51.4]	50 [21.1]	1315 [-15.2]	258 [23.3]	4382 [-2.5]
V-22	7215 [-6.1]	238 [-77.2]	55 [32.6]	1203 [-22.4]	262 [25.0]	4472 [-0.5]

[ ] : Percentage difference with respect to current threshold policy (Vol-28)

* : Including vehicles waiting outside system, i.e. waiting outside the upstream boundary of the freeway and entrance ramps

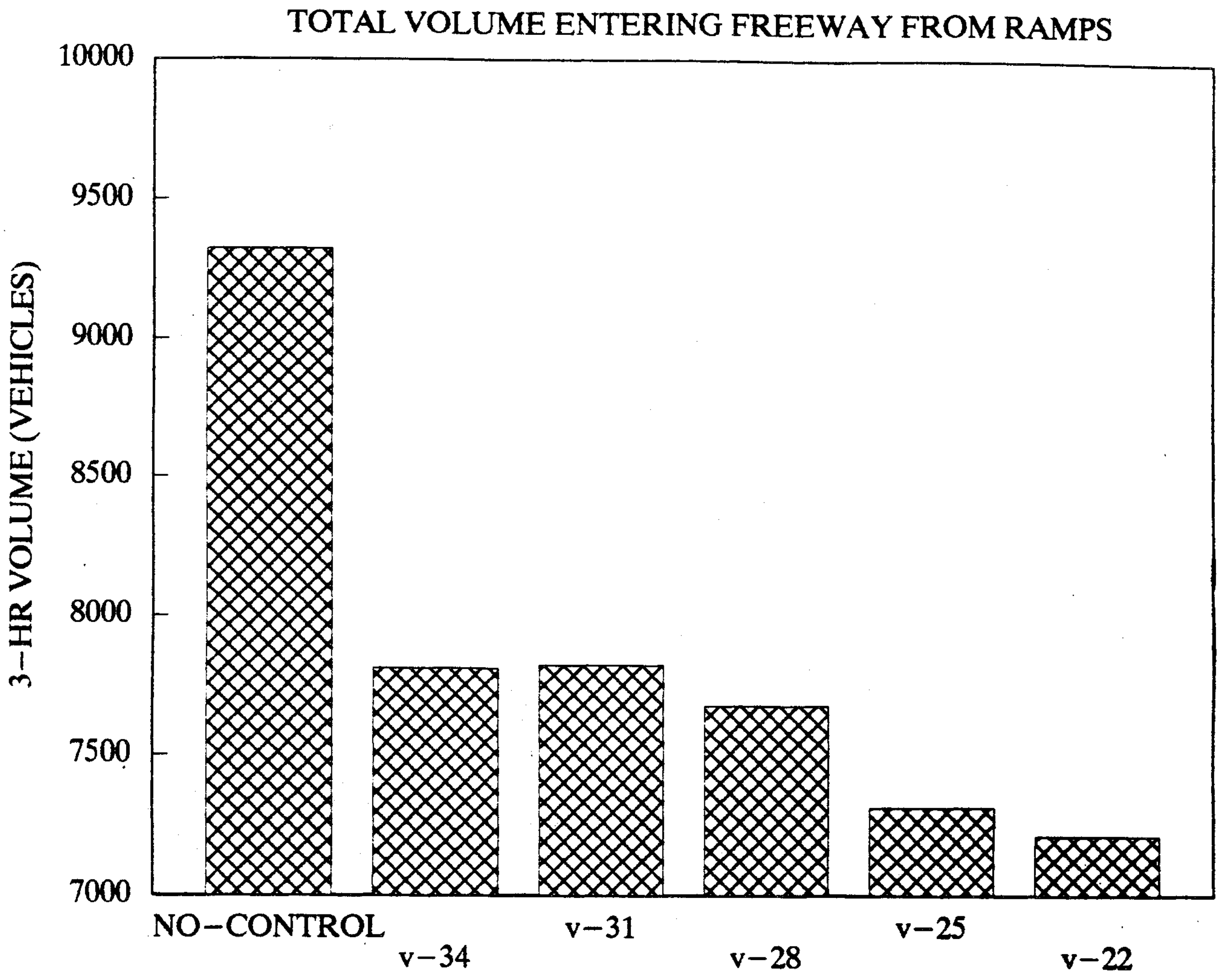

Figure 30. Total Volume Entering Freeway from Ramps for Volume-Threshold Changes

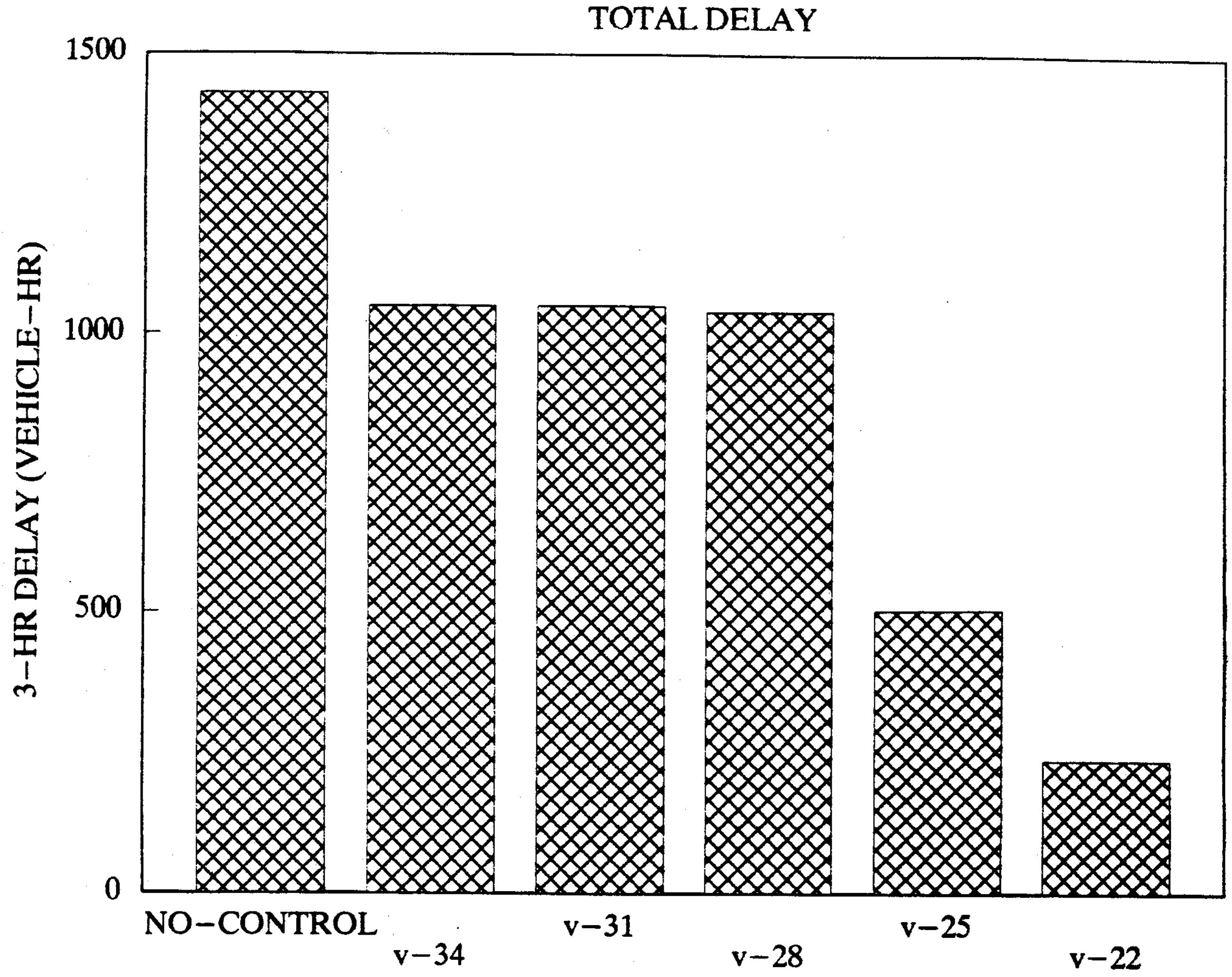

Figure 31. Total Delay for Volume-Threshold Changes

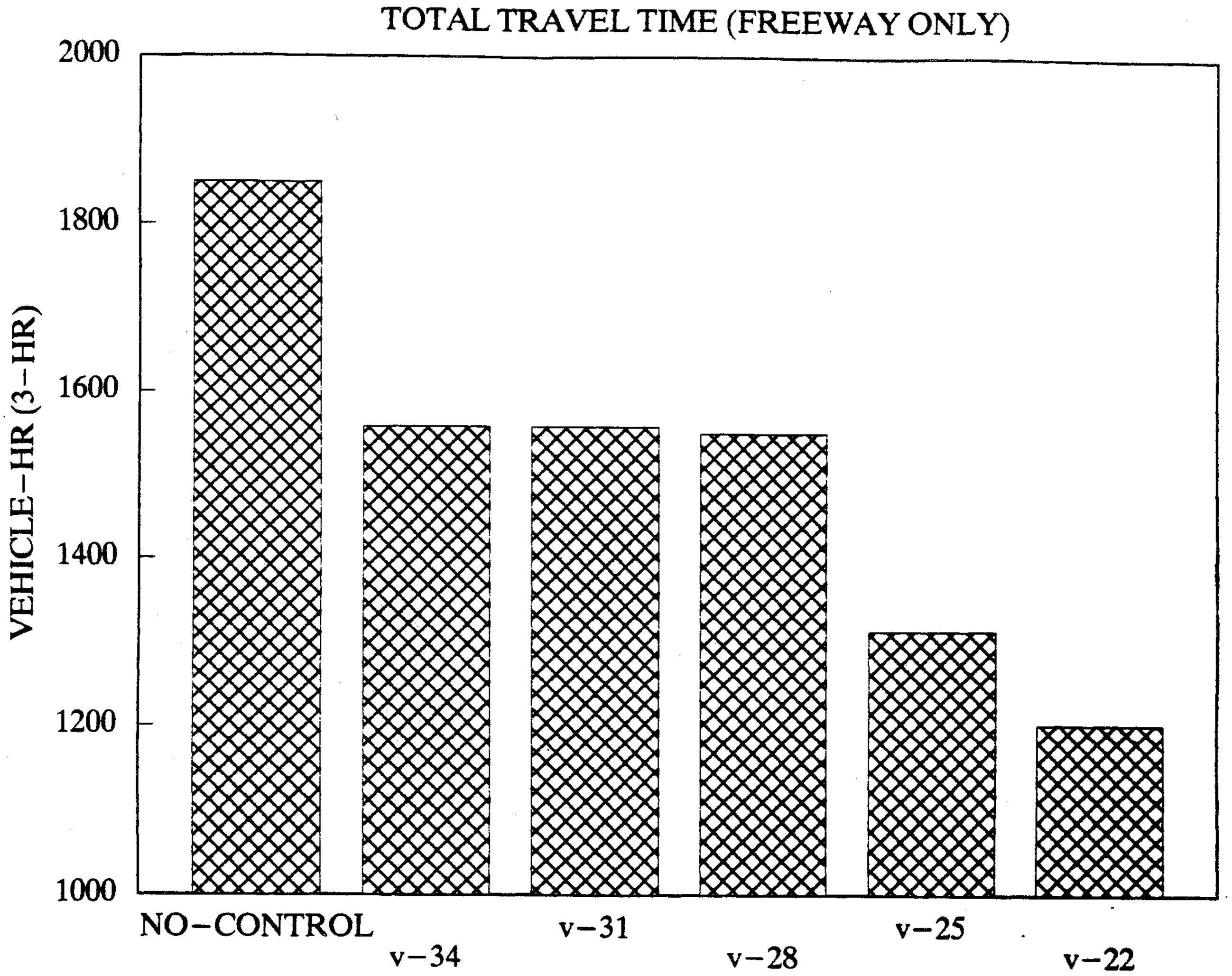

Figure 32. Total Travel Time (Freeway Only) for Volume-Threshold Changes

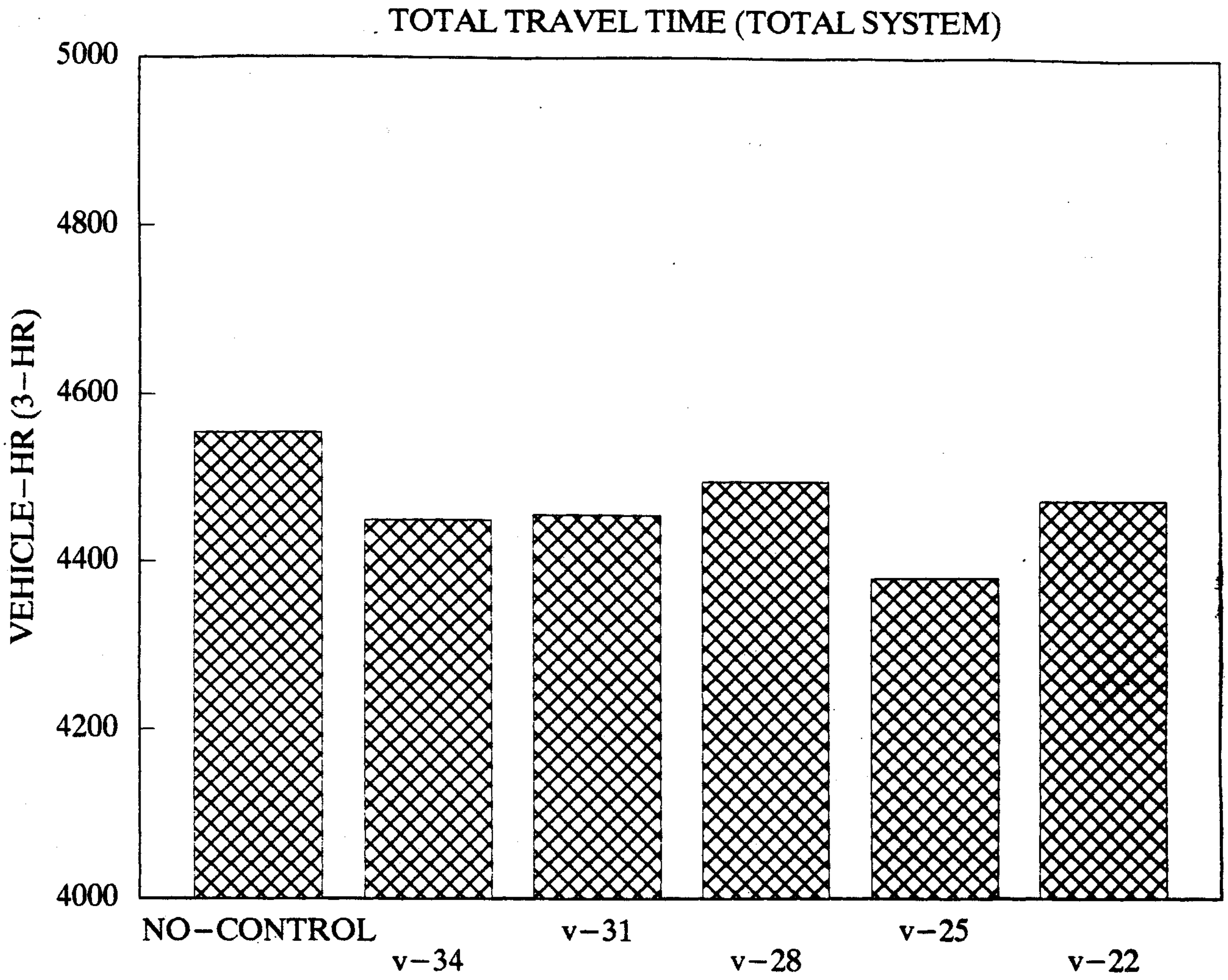

Figure 33. Total Travel Time (Total System) for Volume-Threshold Changes

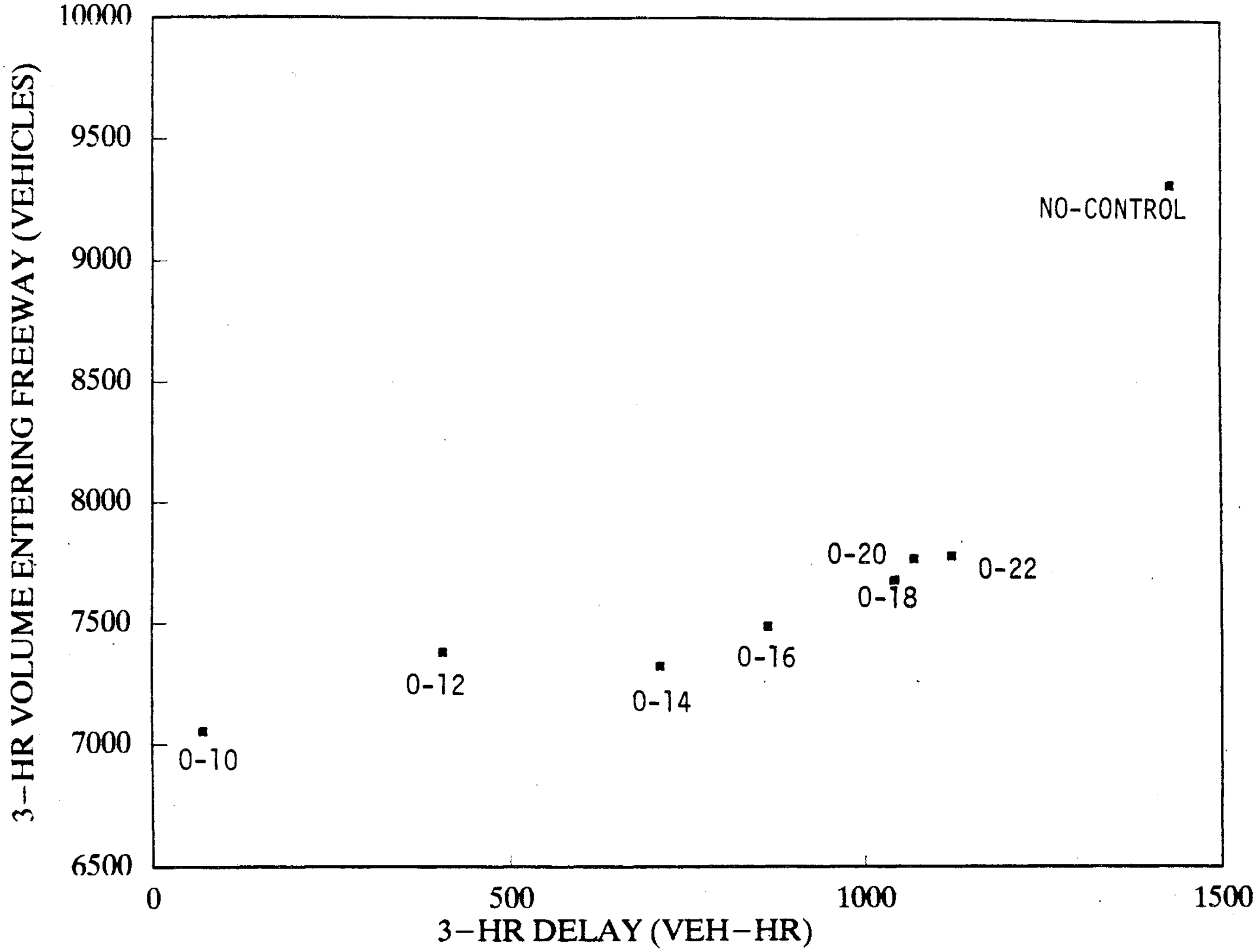

Figure 34. Delay-Volume Space for Occupancy-Threshold Changes

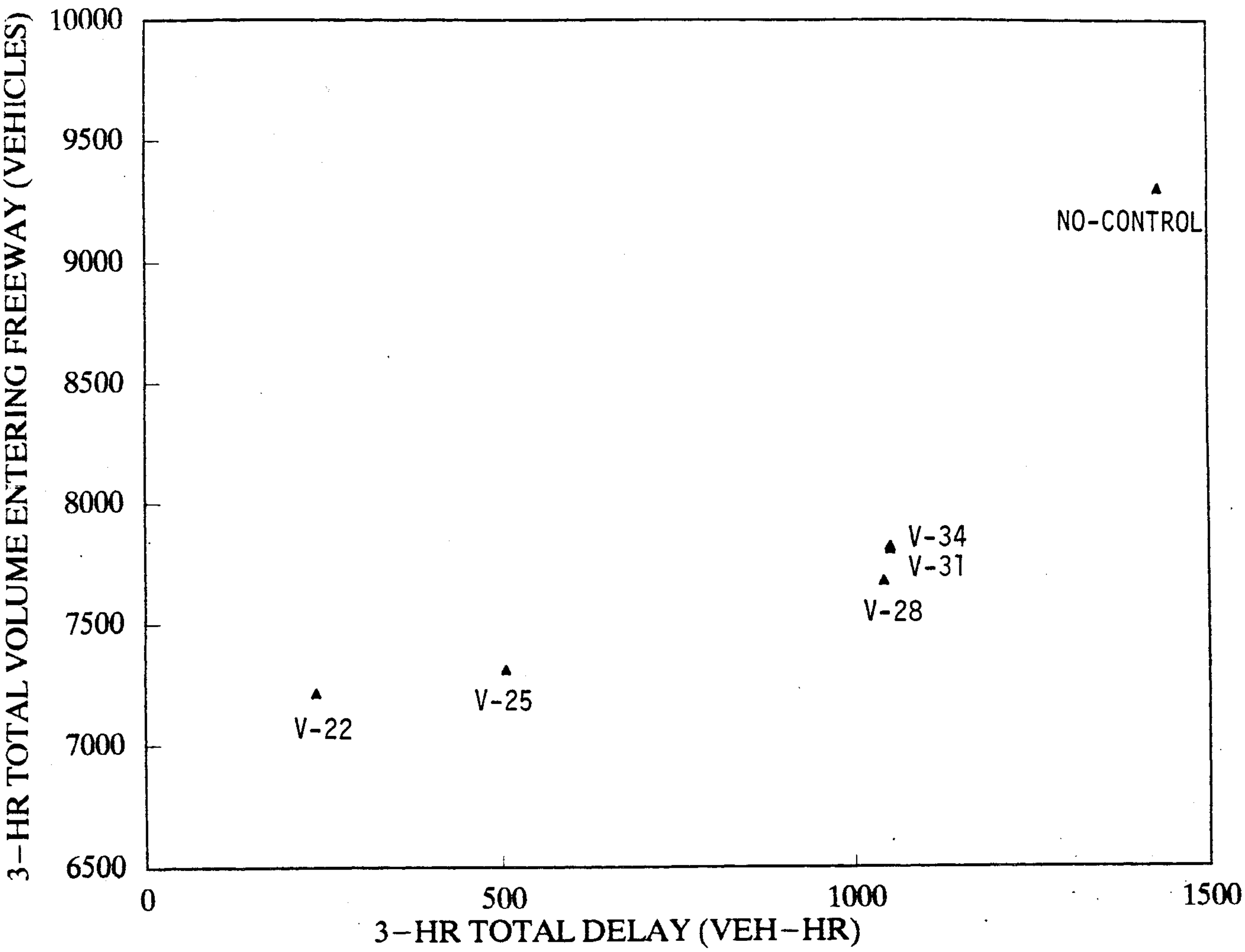

Figure 35. Delay-Volume Space for Volume-Threshold Changes

VI. CONCLUSIONS AND FUTURE RESEARCH NEEDS

Freeway ramp metering has been rapidly expanding in most urban areas in the United States. In the Twin Cities, 122 ramp meters are in operation. Of these, 77 meters are centrally controlled with an automatic rate-selection procedure, and 45 are isolated meters. Further, 82 isolated meters are being installed, and approximately 500 new meters are expected by 1993. Current plans indicate that all isolated meters will be under central control within the next 5 years.

While ramp metering is considered as the most effective form of control in reducing freeway congestion, the state of the art in ramp metering still cannot reflect network-wide traffic conditions and flow interactions between freeways. To be sure, as reviewed in this work, most operational traffic-responsive metering systems set their metering rates based on local congestion levels usually represented by occupancy measurements. While a few systems attempt to reflect system-wide traffic conditions by making adjustments that are based on downstream conditions, existing metering algorithms tend to be reactive rather than predictive. More specifically, they use previously collected data, generally 1-minute old, for calculating metering rates, thereby reacting to freeway conditions rather than preventing congestion. Further, in the automatic rate-selection systems, the updating procedures for thresholds and rate-tables are not always well defined and mostly depend on past experience.

In this research, a method is developed for evaluating traffic-responsive ramp metering strategies and improving freeway performance. The method emulates real-time metering logic and rigorously traces the interactions between automatic rate-selection metering strategies and freeway performance through time. Given a demand pattern and freeway geometrics, it provides assessment of metering strategies that change continuously at very short time intervals, and explicitly treats time delays that can be caused by hardware or introduced by the traffic engineer.

The method was tested with real data in emulating volume and occupancy thresholds, rate tables and automatic rate-selection control strategies on I-35W in Minneapolis. The tests indicated mean percentage errors in the range from 5 to 12 percent. The method was also tested in assessing and selecting ramp control strategies based on freeway performance measures such as total freeway volume and delay. Evaluation results were in the expected direction, and indicated the ability to distinguish small performance differences across strategies.

The method can be a practical tool for facilitating traffic management by aiding in the design of new traffic-responsive strategies at individual ramps or sets of ramps, and in the evaluation of existing control systems and selection of the best metering strategies. Using this tool for considering the trade-offs between performance indicators such as volume increase and delay reduction, a desirable threshold policy can be determined for a given freeway section prior to implementation.

Although the new method can be used to improve the performance of conventional systems, it is still restricted by modeling and hardware limitations. For achieving an on-line performance needed in determining optimal metering strategy configurations in real-time large freeway networks, the following research directions will be addressed in the near future:

1) Extension of the control-emulation method to freeway corridors and large freeway networks.
2) Development of predictive algorithms that can be used in combination with existing automatic rate-selection systems.
3) Development of an efficient optimal control structure, that is computationally feasible and leads to robust control solutions, for real-time management of freeway networks.
4) Development of an efficient on-line optimization algorithm that can be implemented for real time management of large freeway networks using parallel processing techniques.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Rich Lau, Traffic Management Center, Department of Transportation, Minnesota, for his cooperation and valuable suggestions throughout this project, and to acknowledge the Traffic Management Center for providing freeway geometry information and traffic flow data. The Center for Transportation Studies, Department of Civil and Mineral Engineering, University of Minnesota is acknowledged for its support.

REFERENCES

Blinkin, M. (1976). Problem of optimal control of traffic flow on highways. Automation and Remote Control, Vol. 37, No. 5, Part 1, pp. 662-667.

Buijn, H., and F. Middelham (1990). Ramp metering control in the Netherlands. Proceedings of IEE 3rd International Conference on Road Traffic Control, pp. 199-203, London, UK. May.

Cremer, M. (1976). A new scheme for traffic flow estimation and control. In Control in Transportation Systems, Proceedings of the IFAC/IFIP/IFORS 3rd International Symposium, pp. 29-37, Columbus, Ohio, August.

Fox, M. (1990). Ramp metering algorithm description. Ministry of Transportation, Ontario. September.

Georgadakos, G., R. Plum, and P. G. Michalopoulos (1988). KRONOS-5: an interactive freeway simulation program for personal computers. Final Report and User's Manual, Center for Transportation Studies, Department of Civil and Mineral Engineering, University of Minnesota, Minnesota.

Goldstein, N., and K. Kumar (1982). A decentralized control strategy for freeway regulation. Transportation Research B, Vol. 16B, No. 4, pp. 279-290.

Hadj-Salem, H., J. Blosseville, and M. Papageorgiou (1990). ALINEA: A local feedback control law for on-ramp metering; a real-life study. Proc of IEE 3rd International Conference on Road Traffic Control, pp 194-198, London, UK, May.

Hall, C. (1990). Ramp metering algorithm description. Virginia Department of Transportation

Imada, T., and A. May (1985). FREQ8PE : A freeway corridor simulation and ramp metering optimization model. Research Report UCB-ITS-RR-85-10, Institute of Transportation Studies, University of California, Berkeley.

Isaksen, L., and H. Payne (1973). Suboptimal control of linear systems by augmentation with application to freeway traffic regulation. IEEE Transactions on Automatic Control, Vol. AC-18, No. 3, pp. 210-219.

Jacobson, L., K. Henry, and O. Mehyar (1989). Real-time metering algorithm for centralized control. Transportation Research Record, No. 1232, pp.17-26.

Kaya, A. (1971). On the optimization of traffic flow for urban transportation systems. Joint Automatic Control Conference of the American Automatic Control Council, pp. 884-892, St. Louis, Missouri, August.

Keen, K., M. Schofield, and G. Hay (1986). Ramp metering access control in M6 motorway. Proceedings of IEE 2nd International Conference on Road Traffic Control, pp 39-42, London, UK. April.

Kometani, E., T. Hasegawa, and S. Inada (1974). On the operation of the traffic control system on the Hanshin expressway. In Traffic Control and Transportation Systems, Proceedings of the 2nd IFAC/IFIP/IFORS Symposium, pp. 237-248, Monte Carlo. September.

Kostyniuk, L., T. Maleck, W. Tayloret, and A. Hamad (1988). An evaluation of the Detroit freeway operations (SCANDI) project. Michigan State University, East Lansing, Michigan. June.

Lau, R., Traffic Management Center, Minnesota Department of Transportation (1990). Interview. September.

Lipp, L., L. Corcoran, and G. Hickman (1991). The benefits of central computer control for the Denver ramp metering system. Presented at the 70th Transportation Research Board annual meeting, Washington, D.C., January.

Looze, D., P. Houpt, N. Sandell, Jr., and M. Athans (1978). On decentralized estimation and control with application to freeway ramp metering. IEEE Transactions on Automatic Control, Vol. AC-23, No. 2, pp. 268-275.

May, A. (1976). A proposed dynamic freeway control system hierarchy. In Control in Transportation Systems, Proceedings of the IFAC/IFIP/IFORS 3rd International Symposium, pp. 1-12, Columbus, Ohio, August.

McDermott, J., S. Kolenko, and R. Wojcik (1979). Chicago area expressway surveillance and control, Final report. Chicago.

Michalopoulos, P. G. (1984). Dynamic freeway simulation program for personal computers. Transportation Research Record, No. 971, pp. 68-79.

Ovaici, K., R. Teal, J. Ray, and A. May (1975). Developing freeway priority-entry-control strategies. Transportation Research Record, No. 533, pp. 122-137.

Papageorgiou, M. (1980). A new approach to time-of-day control based on a dynamic freeway traffic model. Transportation Research B, Vol. 14B, pp. 349-360.

Papageorgiou, M. (1981). Optimal control of freeway traffic by decomposition methods. Proceedings of International Conference on Control and Its Applications, pp. 369-373, University of Warwick, UK.

Papageorgiou, M., and R. Mayr (1982). Optimal decomposition methods applied to motorway traffic control. International Journal of Control, Vol. 35, No. 2, pp. 269-280.

Papageorgiou, M. (1983). A hierarchical control system for freeway traffic. Transportation Research B, Vol. 17B, No. 3, pp. 251-261.

Payne, H. (1971). Models of freeway traffic and control. Simulation Council Proceedings, Mathematics of Public System, Vol. 1, No. 1, pp. 51-61.

Payne, H., W. Thompson, and L. Isaksen (1973). Design of a traffic-responsive control system for a Los Angeles freeway. IEEE Transactions on System, Man, and Cybernetics, Vol. SMC-3, No.3, pp. 213-224.

Payne H., D. Brown, and J. Todd (1985). Demand-responsive strategies for interconnected freeway ramp control systems, Volume 1: metering strategies. FHWA/RD-85/109, VERAC Incorporated, San Diego, California.

Raffat, A., City of San Diego (1990). Interview. September.

Regional Transit Board (1986). Transit service needs assessment, Phase I report, St. Paul, Minnesota.

Sasaki, T. (1968). Probabilistic models for trip distribution. Proceedings of 4th International Symposium on the Theory of Traffic Flow, pp. 205-210, Karlsruhe, Germany. June.

Stephanedes, Y., and K. Chang (1991). Optimal ramp-metering control for freeway corridors. Proceedings of the 2nd International Conference on Applications of Advanced Technologies in Transportation Engineering, Minneapolis, Minnesota, pp. 172-176. August.

Stephanedes, Y., E. Kwon, and K. Chang (1992). Control emulation method for evaluating and improving traffic-responsive ramp metering strategies. Presented at the 71st Transportation Research Board annual meeting, Washington, D.C., January.

Stevens, M., City of Los Angeles (1990). Interview. September.

Wang, J., and A. May (1973). Computer model for optimal freeway on-ramp control. Highway Research Record, No. 469, pp. 16-25.

Warren, K., Wisconsin Department of Transportation (1990). Interview. September.

Wattleworth, J., and D. Berry (1965). Peak period control of a freeway system-some theoretical investigations. Highway Research Record, No. 89, pp. 1-25.

Yuan, L., and J. Kreer (1968). An optimal control algorithm for ramp metering of urban freeways. Proceedings of 6th annual Allerton Conference on Circuit and System Theory, pp. 52-59, Allerton House, Monticello, Illinois. October.

Yuan, L. and J. Kreer (1971). Adjustment of freeway ramp metering rates to balance entrance ramp queues. Transportation Research, Vol. 5, pp. 127-133.

Zobe, P., W. Johnson, J. Dematte, and W. Cale (1982). Integrated motorist information system (IMIS) Phase III, FHWA/RD-82/108. U.S. Department of Transportation.