Separator

1295-3

A SYNTHESIS OF ADVANCED RAMP METERING STRATEGIES

by

Carroll J. Messer, P.E.
Research Engineer

Joe VanArendonk
Graduate Research Assistant

and

Venugopal Neerudu
Graduate Research Assistant

Research Study Number 0-1295

January 24, 1996

Texas Transportation Institute
The Texas A&M University System
College Station, Texas 77843-3135

TABLE OF CONTENTS

1. INTRODUCTION TO RAMP METERING

1.1 BACKGROUND

The TxDOT has been upgrading its urban freeway system over the past decade in all major metropolitan areas of the state to increase capacity and improve operations. This added freeway capacity has permitted the economy to continue to expand to the general benefit of all of its citizens. However, the prospect of another round of freeway capacity upgrades in urban areas is dim. Consequently, as traffic volumes have continued to grow at about 3% per year, traffic congestion and air quality problems are again becoming severe in some locations as evidenced by the fact that several Texas cities have been designated "non-attainment" areas under current federal ozone standards. Adding new freeway lanes in cities like Houston, designated a "serious non-attainment" area, may not be permitted under CAA 1990. Ramp metering is considered a congestion mitigation strategy for air quality improvement (CMAQ) within ISTEA 1991, together with other freeway traffic management (FTM) strategies.

Two decades have past since any large-scale ramp metering has operated in Texas *(1)* so TxDOT initiated this research project to address some additional operational issues identified beyond those addressed in initial efforts and even recent work underway within Project 1232, the "Urban Highway Operations Research and Implementation Program." Task 5 of Project 1232 basically developed a microcomputer-based support system (Rambo) to expedite the implementation of TxDOT's existing ramp metering specification, TxDOT-550-80-950-01. Subsequent work conducted within Project 1232 during 1995 provided an enhanced special "Ft. Worth" specification, dated July 31, 1995, which further relies on the Rambo support system and general methodology. Implementation of ramp metering on five ramps along Texas 360 in Arlington is expected to be underway by Fall 1996. A similar schedule of ramp metering is planned for US 290, Austin Freeway, in northwest Houston, within the context of an ITS Corridor Demonstration effort during 1996. Thus, the prospects of sizeable ramp metering systems in Texas merits the continued investigation, synthesis and assessment of ramp metering technology being developed by this project and by other operating/research agencies.

1.2 OBJECTIVES OF PAPER

This paper will present the most recent operational results from ramp metering systems employed in New York and the latest ramp metering strategies described in the literature being developed for possible deployment at other major freeway traffic management projects, namely, at New York, Seattle, and Anaheim (Los Angeles). An update on the use of Federal Highway Administration's FREESIM microscopic simulation program to ramp metering is also provided. FREESIM's limitations have been recognized by the researchers in Seattle. As a result, it is believed that FHWA is currently working on some suggested ramp metering enhancements to FREESIM.

Another purpose of this paper is to serve as a focus for discussion with TxDOT FTM managers of optional ramp metering technologies that might be used in Texas beyond those contained in the current specifications noted above for Ft. Worth. In addition, the collection of possible technologies described will serve as a guide to assist in the selection of future strategies that might be selected, implemented, and evaluated within this research program.

1.3 LITERATURE RESEARCH METHODOLOGY

The research team conducted a literature search of on-line sources and examined numerous documents on ramp metering. Those metering concepts that seemed to offer new features, contained new technology engines, and appeared to be supported by a State Department of Transportation were given the most consideration. Most of the descriptions, technology and characterization of the following ramp metering strategies were provided by the original authors of the papers cited, and no credit is claimed by the research team for any of the referenced materials surveyed. We elected to extract much of the materials as written so that a clearer understanding of the fundamentals described by the cited research would be provided to our FTM panel members. Any false citations or misrepresentations of the following ramp metering strategies are surely unintentional errors of the 1295 research team. Any assessments provided by the research team are offered in the context of possible applications in Texas within the near future, and does not necessarily reflect the technical merits of the strategy noted.

2. NEW YORK'S INFORM RAMP METERING SYSTEM

2.1 INTRODUCTION

The ramp metering system in Long Island, New York is part of a larger traffic management system know as INFORM (Information FOR Motorists, formerly known as the Integrated Motorist Information System-IMIS). This system has many of the same features as the proposed Houston Freeway Traffic Management (FTM) system. Both systems include integrated electronic traffic monitoring, variable message signing, ramp metering, and related strategies to optimize traffic flow through a heavily congested corridor. More specifically, the proposed ramp metering systems in Texas can benefit from the lessons learned from the INFORM system.

INFORM adopted an unusual "tolerable" metering policy (in contrast to demand/capacity) which allowed meters to turn off or "queue off" as the ramp's queue detector became highly occupied. "Queue off" periods could last for as much as half the operable time of the meter, thereby reducing ramp queues and delay times. This unusual permissive metering system allowed the meter to adapt to existing conditions by being restrictive at low volumes yet more tolerable to local traffic during high volumes. Delays, therefore, were perceived by local motorists to be less than one minute on the average and opinions of the system in general were generally favorable.

The proposed Houston/Ft. Worth FTM systems may benefit from a similar "tolerable" system. Without being able to reliably predict the public reaction and traffic conditions at the implementation stage, of ramp metering, it may be wise to employ a similar system. Such a system could meter most traffic conditions yet not create too much regulation on traffic flow.

Similar to some Houston ramp meters, the INFORM meters had been standing dormant for two years prior to being operated. This necessitated a strong public relations campaign to inform drivers that a new ramp control strategy would be employed. Local politicians and the media were involved as well as several police agencies. To this date, motorist compliance rates have been commendable and the system has become an integral part of the INFORM system.

The following is a summary of the ramp metering status in a report written on "INFORM" by Steven A. Smith of JHK & Associates in January 1992 *(2)*. Ramp metering has been an integral part of the INFORM system from its original conception. The first ramp meters became operational in December of 1989 and by March of 1990 more than 40 were working. There are currently 70 ramp meters in operation, 50 on the Long Island Expressway (LIE) and 20 on the Northern State Parkway (NSP)*(3)*. The percent of entering traffic metered by roadway and direction is:

- LIE westbound a.m. metering - 36.5 percent;
- LIE eastbound p.m. metering - 50.9 percent;
- NSP/GCP westbound a.m. metering - 20.6 percent; and
- NSP/GCP eastbound p.m. metering - 16.1 percent.

Some ramp volumes changed considerably between the time of design and 1990. Several major office parks were developed during that period which have added many vehicles to the peak hour volume. It became very difficult to maintain metering operation at these high-volume locations.

Each installation contains a ramp meter, an advanced warning sign activated prior to the meter turn-on, signs at the ramp meter location stating "STOP HERE ON RED" and "ONE VEHICLE PER GREEN," input and output detectors, and a queue detector. Input/output detectors control the release of vehicles entering on the ramp and also provide volume and occupancy information to the central computer. The queue detector measures occupancy only and serves to warn the central computer of queues backing into the arterial from the on-ramp traffic.

2.2 METERING OPERATIONS

Modes of Operation

Ramp meters can be operated in manual, time of day, or traffic responsive modes. These operating modes are described below:

- *Manual operations* - system operators may select any individual or group of ramp meters to operate at a specific metering rate. Meters can be turned on and off, or the metering rate changed by the operator in the control center. This mode has been primarily used during the testing stage of ramp metering when each meter was being brought on line.

- *Time of day mode* - In time of day mode, the operator specifies individual turn-on time, turn-off time and metering rate for each individual ramp. The system then initiates operation and ceases operation at specified times. There is no opportunity for varying time-of-day metering rate within a given metering period (turn-on time to turn-off time). Metering rates would have to be changed by turning metering off, then turning metering back on again. Time-of-day was the primary mode of operation through April 1990. Time-of-day mode also contains a provision for automatic metering shut-off and turn-on in the event that a queue extends back to a queue detector. An occupancy threshold is established individually for each metering location, both for the threshold over which metering would be shut off, and under which metering would be turned back on. Typical turn-off occupancy thresholds have been between 15 and 25 percent. The turn-on threshold is only used for reinitiating ramp metering after it has been turned off within a metering period. It is not used to sense when metering should be initially turned on.

- *Traffic responsive mode* - Traffic responsive mode adjusts the metering rate in response to mainline and ramp traffic conditions. The traffic responsive mode adds an additional dimension to the management of queues on metered ramps. As the occupancy of the queue detector increases, the traffic responsive metering algorithm increases the metering rate (limited to the maximum metering rate) to avoid or forestall a shutdown of the metering operation on that ramp. On the mainline, the traffic-responsive algorithm examines both the upstream and downstream detector stations. Degradation of speed on the mainline will result in a reduction of the metering rate. However, this action will be overridden by excessive queuing on the ramp itself. Thus, the entire metering operation is ultimately controlled by the ability of the ramps to store traffic.

The minimum and maximum metering rates used for any mode are 300 VPH and 800 VPH, respectively. A maximum rate of 900 VPH was originally planned; however, it was determined that this rate provided insufficient red time for a driver to come to a complete stop. Two-lane metering has been proposed as an experimental design; however, to this date no two-lane design has been implemented in the INFORM project.

Implementation Plan

The ramp metering system was turned on in stages. An implementation plan was prepared by the operations contractor in April 1988. The implementation plan identified groupings of ramps according to their status of readiness and then rated them in order to prioritize the implementation. The initial turn-on of a ramp meter took between one day and one week of careful observation in the field. Temporary signing was installed in advance of the metering date, identifying the date on which the metering was to be initiated at the site. This was particularly important in light of how long the meters had been visible (more than two years) without being operated.

Ramp metering implementation took place over a period of approximately one year, beginning with the first ramp turn-on in December of 1988. This was a longer period than first envisioned, due to hardware problems and to modifications believed necessary to provide for safe operation.

Implementation was preceded by an extensive public relations campaign that officially began on December 13, 1988 with a media event in the control center. In the campaign, the term "merge light" was used in place of "ramp meter" as it was believed by the public relations consultant that the new term conveyed a more acceptable message to the public and was easier for them to understand. Information was also conveyed through monthly incident management meetings, which were attended by most of the affected police agencies in the INFORM corridor. The operations staff had to be sensitive to the safety-related concerns of the police. A policy has remained in force that if a patrolman requests a specific ramp meter to be shut down, regardless of the reason, the operations staff will shut it down without question. Thus, the police have ultimate control over the ramp metering operation. However, such requests have rarely been generated by the police.

In the initial three months of ramp metering operation, there were six minor rear-end collisions on the metered on-ramps. In order to reduce the number of accidents, experimentation began with using a high-intensity strobe light in a ring on the red signal head. This device called the drivers' attention to the ramp meter. Since it proved to be successful, they have been installed on all ramp meters within the system.

2.3 METERING EFFECTIVENESS

Traffic Performance

A series of operational tests were conducted to assess the operational performance of the various modes of operation. Metering in March 1990 was conducted in the time-of-day mode, while the periods in April/May and June 1990 were conducted in the traffic responsive mode. In both modes, there was the possibility that excessive queues would force the shut-off of metering. As noted previously, this is referred to as the "queued-off" condition. The amount of time that metering stays on, in combination with the metering rate, defines the degree of restriction in the metering plan.

A summary of ramp metering operations at 40 ramps, presented in Table 2.1, shows that the ramps with heavier volumes were the ones that typically queue off more frequently. A number of ramps were queued off for nearly half of their 2-hour targeted metering period. This points to the difficulty of sustaining ramp metering under high volume conditions, particularly if only single lane metering is available.

There were significant periods when ramp metering was shut down to avoid surface street impacts. Field observation of some surface streets impacts indicated that the queuing impacts were a very real concern to motorists and that the decision to continue to meter restrictively at the moment would create major surface street traffic problems. The expected impacts and subsequent public outcry were the major incentives given by INFORM operations staff for maintaining the "tolerable" ramp metering policy.

Table 2.1 Summary of ramp metering operation

RAMP NO	RAMP NAME	TYPE OF METERING	TIME ON	AVG. MINUTES ON	AVG. TIMES QUE'D OFF	AVG. OFF TIME	BASE METERING RATE	PEAK 15_MIN. VOLUME	AVG. QUEUE
LIE EB RAMPS									
1	MAIN ST	TOD	1600	114.8	0.2	26	800	618	
1	MAIN ST	AUTOMATED	1600	120	0	0		618	1.2
2	161ST ST.	TOD	1600	113.2	0.4	17	800	642	
2	161ST ST	AUTOMATED	1600	120	0	0		642	
3	UTOPIA PKWY	TOD	1600	120	0	0	800	706	
3	UTOPIA PKWY	AUTOMATED	1600	118.8	0.1	12		706	4.5
4	OCEANIA ST.	TOD	1600	120	0	0	800	361	
4	OCEANIA ST	AUTOMATED	1600	120	0	0		361	0.6
5	SPRINGFIELD	TOD	1600	120	0	0	500	296	
5	SPRINGFIELD	AUTOMATED	1600	120	0	0		296	0.6
6	L. NECK PKWY	TOD	1600	110	0.2	50	800	583	
6	L. NECK PKWY	AUTOMATED	1600	117	0.1	27		583	1.3
7	COMMUNITY DR	TOD	1600	115.8	0.25	97	800	987	
7	COMMUNITY DR	AUTOMATED	1600	134.1	1.4	31.4		987	
8	NEW HYDE PK	TOD	1600	93.8	0	105	800	715	
8	NEW HYDE PK	AUTOMATED	1600	67.6	1.3	39.3		715	
9	SEARINGTOWN	TOD	1600	120	0	0	800	808	
9	SEARINGTOWN	AUTOMATED	1600	108	0.4	30		808	3.9
10	WILLIS AVE	TOD	1600	120	0	0	800	594	
10	WILLIS AVE	AUTOMATED	1600	120	0	0		594	2.3
11	GLEN COVE RD	TOD	1600	120	0	0	800	461	
11	GLEN COVE RD	AUTOMATED	1600	120	0	0		461	1.6
13	S.O.BAY RD	TOD	1600	55.7	1	64.3	800	903	
13	S.O.BAY RD	AUTOMATED	1600	88.4	1.8	17.6		903	3.4
15	RT 110 NORTH	TOD	1600	120	0	0	800	727	
15	RT 110-NORTH	AUTOMATED	1600	120	0	0		727	5.8
16	PINELAWN RD	TOD	1600	114.5	1	11	800	1022	
16	PINELAWN RD	AUTOMATED	1600	100.1	1.3	17.3		1022	11.1
21	VANDERBILT	AUTOMATED	1600	90.5	0	0		1469	0.7
22	RT 111	TOD	1600	88	3	10.7	800	832	
22	RT 111	AUTOMATED	1600	73.7	4.1	11		832	7.4

Table 2.1 Summary of ramp metering operation (continued)

RAMP NO	RAMP NAME	TYPE OF METERING	TIME ON	AVG. MINUTES ON	AVG. TIMES QUE'D OFF	AVG. OFF TIME	BASE METERING RATE	PEAK 15_MIN. VOLUME	AVG. QUEUE
LIE WB RAMPS									
24	RT 111	TOD	700	60	0	0	800	534	
24	RT 111	AUTOMATED	700	60	0	0		534	
25	VANDERBILT	TOD	700	60	0	0	700	426	
25	VANDERBILT	AUTOMATED	700	57	1.25	6		426	
26	COMMACK RD	AUTOMATED	700	57.2	3.4	19.3		818	
28	BAGATELLE RD	TOD	700	20.9	0	0	800	957	
28	BAGATELLE RD	AUTOMATED	700	13.4	1.9	24.2		957	
30	ROUND SWAMP	TOD	718	50.3	0	0	800	266	
30	ROUND SWAMP	AUTOMATED	700	60	0	0		266	
31	SUNNYSIDE	TOD	732	50.3	0.2	0	800	128	
31	SUNNYSIDE	AUTOMATED	800	58.8	0	0		128	
33	JERICHO TPK	TOD	700	118.6	0	0	800	521	
33	JERICHO TPK	AUTOMATED	700	120	0	0		521	0
34	GLEN COVE RD	TOD	700	103	0	0	800	594	
34	GLEN COVE RD	AUTOMATED	700	95.1	1.3	16.6		594	6.5
36	SEARINGTOWN	TOD	700	118.6	0	0	800	331	
36	SEARINGTOWN	AUTOMATED	700	113.1	0	0		331	0.7
37	SHELTER ROCK	TOD	710	118.3	0	0	800	243	
37	SHELTER ROCK	AUTOMATED	700	120	0	0		243	0.6
38	NEW HYDE PK	TOD	600	109.3	0	0	800	479	
38	NEW HYDE PK	AUTOMATED	631	108.4	0	0		479	0
39	COMMUNITY DR	TOD	700	100.1	0.1	32	800	404	
39	COMMUNITY DR	AUTOMATED	700	120	0	0		404	2.2
40	LAKEVILLE	TOD	700	106.6	0	0	800	313	
40	LAKEVILLE	AUTOMATED	700	120	0	0		313	0.3
41	L. NECK PKWY	TOD	700	114.3	0	0	800	547	
41	L. NECK PKWY	AUTOMATED	700	120	0	0		547	5.6
44	UTOPIA PKWY	TOD	810	101.5	0	0	800	999	
44	UTOPIA PKWY	AUTOMATED	702	105	0.1	22		999	1
46	MAIN ST	TOD	700	117.6	0	0	500	239	
46	MAIN ST	AUTOMATED	700	122	0	0		239	0.5

Table 2.1 Summary of ramp metering operation (continued)

RAMP NO	RAMP NAME	TYPE OF METERING	TIME ON	AVG. MINUTES ON	AVG. TIMES QUE'D OFF	AVG. OFF TIME	BASE METERING RATE	PEAK 15_MIN. VOLUME	AVG. QUEUE
NSP EB RAMPS									
47	MARCUS AVE	TOD	1600	105	1	75	800	1132	
47	MARCUS AVE	AUTOMATED	1600	71	4	25.4		1132	6.1
48	SHELT ROCK N	TOD	1700	60.2	0	0	800	105	
48	SHELT ROCK N	AUTOMATED	1703	45.6	0	0	800	105	0
49	WILLIS AVE	TOD	1731	101.6	0	0	500	326	
49	WILLIS AVE	AUTOMATED	1600	119.5	0	0		326	1.2
50	IU WILLETS	TOD	1600	118.8	0.2	5	600	438	
50	IU WILLETS	AUTOMATED	1600	180	0	0	600	438	
51	POST AVE N	TOD	1600	3	0	0	800	1223	
51	POST AVE N	AUTOMATED	1601	28.6	2.8	47.6		1223	1.1
52	RTE 106 NB	TOD	1600	119.2	0.2	3	800	416	
52	RTE 106 NB	AUTOMATED	1600	99.3	0.8	4.3		416	0.8
53	S.O.BAY RD N	TOD	1600	119.5	0.3	2	600	277	
53	S.O.BAY RD N	AUTOMATED	1600	119.4	0.1	4		277	1.4
54	RTE 110 NB	TOD	1712	87	0.3	12	800	677	
54	RTE 110 NB	AUTOMATED	1600	102.8	1.8	9.7		677	8.9
NSP WB RAMPS									
55	RTE 110 SB	TOD	700	112.7	1	7.3	NA	604	
55	RTE 110 SB	AUTOMATED	700	100.2	2.1	10.1		604	
56	S.O.BAY RD S	TOD	700	44.9	0.1	0	800	89	
58	ROSLYN RD	AUTOMATED	700	67.8	2.3	22.4		886	
59	WILLIS AVE	TOD	710	96.5	2.2	10.8	700	619	
59	WILLIS AVE	AUTOMATED	700	76	4.6	9.6		619	3.6
60	SHELTER ROCK	TOD	600	79.7	2.1	17.3	800	798	
60	SHELTER ROCK	AUTOMATED	600	86.2	2.8	12.1		798	2.5
61	NEW HYDE PK	TOD	700	120	0	0	800	487	
61	NEW HYDE PK	AUTOMATED	700	112.8	1.5	4.8		487	2.3
62	LAKEVILLE SB	TOD	700	120	0	0	800	191	
62	LAKEVILLE SB	AUTOMATED	700	120	0	0		191	0.4
63	L. NECK PKWY	TOD	700	111.9	0.6	8.8	700	606	
63	L. NECK PKWY	AUTOMATED	700	101.8	2.4	7.6		606	4.7
64	UNION TPKE	AUTOMATED	700	121.4	0	0	600	702	
65	FR LEWIS SB	TOD	710	114.6	0.3	6	500	323	
65	FR LEWIS SB	AUTOMATED	700	104.7	0.1	6		323	0.6

Most ramp queues observed were less than five vehicles. This was due to the propensity for the system to shut certain ramp meters down in response to excessive queuing on the ramps. The average queue computation included those periods when the meter was "queued off," sometimes as much as half the operating time, and when there was no queue at the meter.

Freeway Throughput

Due to the minimum ramp delays and "queued off" periods, the freeway throughput volumes were not considerably affected. The a.m. peak period speeds for the March 1990 metering case increased 3- to 8- percent over the March 1990 nonmetering case and 13-percent over the spring 1987 case. The p.m. peak period speeds for the March 1990 metering case were unchanged from the March 1990 nonmetering case and increased 13-percent over the spring 1987 case. VMT increased approximately 1-percent over the March nonmetering case and approximately 5-percent over the spring 1987 case.

To provide a perspective, an improvement in speed of 10-percent would result in approximately three million vehicle hours of delay saved annually. Thus, there is potential for substantial reduction in vehicle hours due to ramp metering.

The maximum increase in throughput in a bottleneck section for the metering scenario was 7-percent. Thus, ramp metering may produce marginal increases in throughput through bottleneck sections, but not likely more than 2- to 3- percent, on average.

Ramp Delay

The number of vehicles hours of ramp delay can be computed by multiplying the average number of vehicles in queue for each ramp by the amount of time that metering was to have been active (usually two hours). The estimated VHT due to ramp delay was 86 vehicle hours for the a.m. metering period for both the (LIE) and Northern State Parkway/Grand Central Parkway (NSP/GCP) and 147 vehicle hours for the p.m. metering period. In each case, the VHT delay represented only about one tenth of 1-percent of the total VHT for the respective peak periods which was an incidental amount of delay to entering traffic.

It should be recognized that the ramp delay was probably less than what it was expected to be with ramp metering. There were low volumes at some ramps, while other ramps experience so much queuing under metering that the meters were queued off, eliminating the ramp delay. The logical conclusion, then, was that queue storage on the ramps is a critical element of system design. The major factors in creation of storage capacity were ramp length, location with respect to nearby surface streets, and two-lane versus single lane metering. While work had been done on the ramp metering algorithm to reduce the propensity for queued off meters, the ability to meter and the flexibility in metering was seriously compromised by not having two-lane ramp metering capability available. Thus, careful consideration of ramp volumes, storage capacities, and operational policy on queue management is essential in system design.

Motorist Compliance

Data on motorist violations of the ramp metering signals were accumulated by the system based on analysis by the input/output detectors coordinated with timings of the ramp controller. These were field checked for reasonableness. Motorist compliance was good. Percent compliance ranged from a low of 74 percent to a high of 96 percent. The average compliance on the NSP/GCP ramps was 85 percent. The average on the LIE was 83 percent.

Driver questionnaires indicated that drivers understood the legal status of the "merge lights." The vast majority of drivers recognized the merge lights (ramp meters) as a legal traffic control device that must be obeyed. However, a sizeable proportion (over 25 percent) perceived that there would be no penalty if they went through a red merge light. While observed compliance with the ramp meters was quite good, the perception among drivers was that the meters were not backed by significant enforcement power.

A correlation can be made between ramp delay, driver perceptions, and motorist compliance. The high compliance rates were most probably due to the short delay times at the ramp meters, a favorable opinion of ramp meters (see driver perception below), and an understanding of their legal status. See Fig. 2.1 for further summary statistics.

Driver Perceptions

The term "merge light" was used in place of "ramp meter" in the questionnaire, as this was the name for ramp metering signals used in the public relations campaign. The results indicated that approximately one-fifth of all drivers had no opinion about merge lights, and that the remainder was split approximately 50-50 on whether they were a good idea or not a good idea. LIE and NSP/GCP commuter opinions, on the other hand, showed a stronger positive response to merge lights, with 60-percent responding favorably. The most frequently checked responses concerned the negative aspects of ramp metering. Nearly 45 percent of all drivers questioned referred to the creation of backups at the ramps, although, the waits were perceived to be one minute or less by almost 90-percent of the drivers. The difficulty of merging into traffic from a stop was checked by over 40 percent of the drivers. The most frequently listed benefit was that the merge lights can help reduce merge accidents.

Some 15 percent of those encountering a red merge light indicated that they frequently use the service road or another roadway to avoid waiting at the merge lights. Another 27 percent indicated that they do this occasionally. This suggested that ramp metering was, at least in part, having some diversionary effects.

A review of the driver perceptions of ramp metering indicated that there was still a gap in their understanding of the function of ramp metering and their responsibility toward it. Nevertheless, field reviews of the ramp metering operation indicated generally satisfactory operation results.

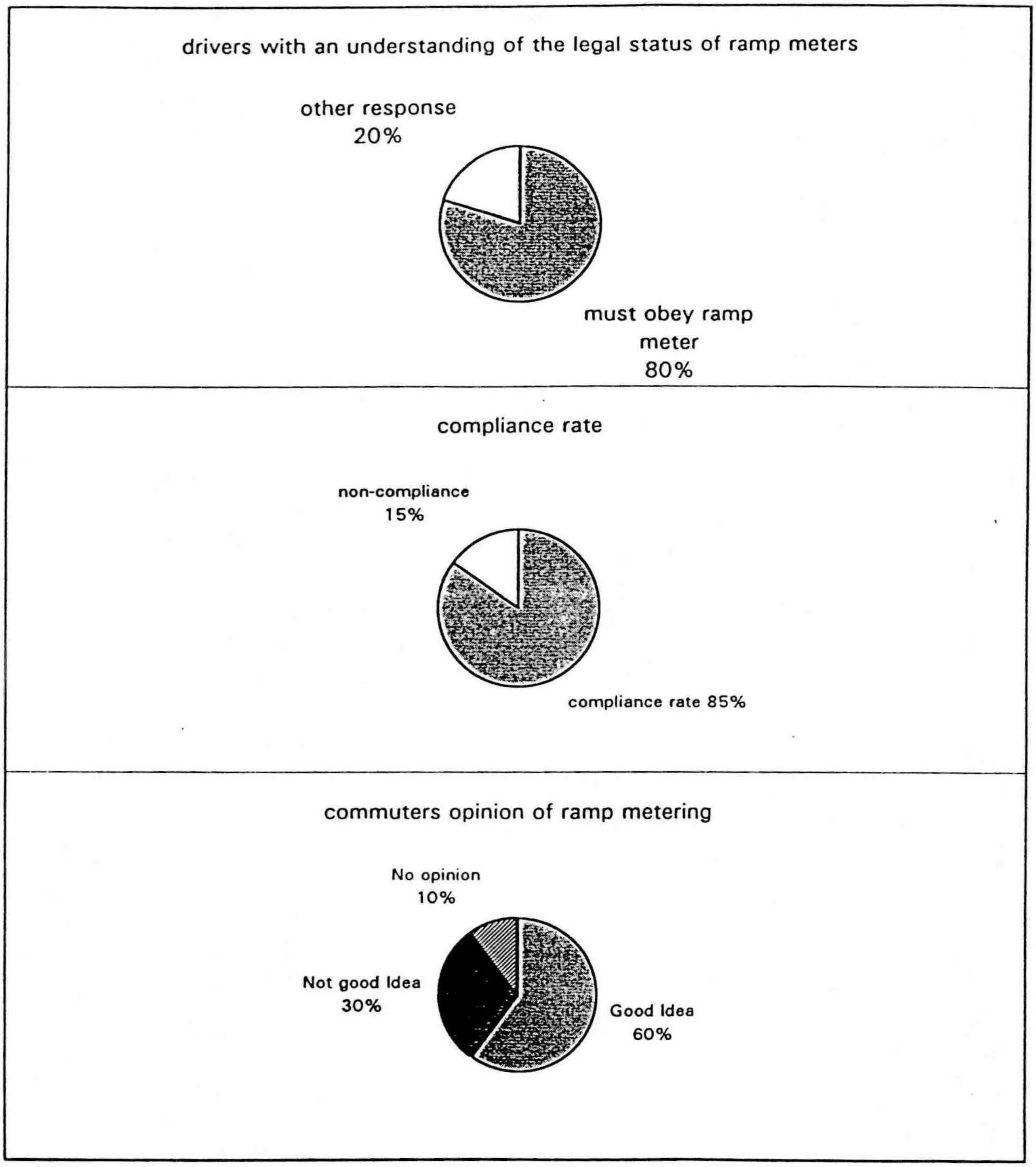

Figure 2.1 Compliance vs. driver perception.

2.4 Conclusions and Recommendations

The study concluded that INFORM is constrained by the limitations in the number of ramps metered, in the storage areas to manage queues, and in the maximum metering rates for single-lane metering. INFORM does not have sufficient ramp metering control over enough traffic to produce a noticeable, sustained change in freeway speeds. Significant use of two-lane metering is needed to exercise greater control over high-volume on-ramps. Additional ramps also need to be metered, including selected freeway-to-freeway ramps before adequate control can be established.

In order to improve ramp metering effectiveness, queuing lanes should extend onto the service roads. Such queuing should be acceptable as long as serious interference with nearby cross streets does not occur and excessive ramp delays do not result. Since the high volume ramps do not have adequate storage capacity, the system relies on the ability of the ramp meters to "queue off." This has had a positive impact on motorist compliance, excess delays, and spillback of queues into arterial streets. To some extent this may have been the reason that the system has received generally favorable reviews and remained in operation.

3. SEATTLE RAMP METERING SYSTEM

Research has been underway in Seattle to continuously improve freeway traffic operations along Interstate 5 where ramp metering is a critical component to the overall freeway traffic management system. A six-year study, during which ramp metering was implemented, indicated that the travel time for a specific 11.1 km section of Interstate 5 decreased from 22 to 11.5 minutes, and the accident rate decreased by 39 percent. During this study, the mainline freeway volumes increased by 86 percent northbound and by 62 percent southbound *(4)*. The Seattle researchers believe that even slight improvements in the ramp metering algorithm may produce significant returns.

3.1 CURRENT METERING ALGORITHM

The ramp metering algorithm currently used in Seattle is relatively sophisticated, but has some limitations *(5,6)*. Features that make Seattle's ramp metering algorithm more sophisticated than others in this country include a volume reduction based on downstream bottlenecks and further local adjustments, such as advanced queue override *(5)*. The current ramp metering system responds to real-time loop detector data through a centralized computer and field-located microprocessors. The controller calculates both a local metering rate and a bottleneck metering rate and uses the more restrictive of these two rates. The local metering rate is based upon adjacent upstream mainline occupancies. Linear interpolation between the actual occupancy and predetermined metering rates for given occupancies determines the local metering rate.

The bottleneck algorithm is activated when the following two criteria are met: 1) a downstream bottleneck-prone section surpasses a predetermined occupancy threshold, and 2) that section stores (is storing) vehicles. A section stores vehicles when more vehicles enter the section than leave the section. When these two criteria are met, the algorithm reduces the number of vehicles entering the freeway by the number of vehicles being stored in the bottleneck section. This volume reduction is distributed over the upstream ramps that can influence that bottleneck. The number of ramps that can affect a bottleneck varies for each site. A weighting factor for each ramp determines the fraction of the volume reduction targeted for that ramp.

After selecting the more restrictive of the local and bottleneck metering rates, the controller further adjusts the metering rate on the basis of local conditions. Queue adjustments prevent the ramp queue from blocking adjacent arterials. A queue adjustment occurs when the occupancy on a ramp exceeds a predetermined threshold for at least a specified duration. In this event, the metering rate increases by a certain number of vehicles per minute; the increase is dependent on which of the two occupancy and duration threshold sets has been exceeded. An advance queue adjustment occurs when a loop detector located near the arterial activates over a particular occupancy threshold for at least a specified duration. The advance queue adjustment also increases the metering rate by a specific number of vehicles per minute. High occupancy vehicle (HOV) adjustment accounts for the difference between the number of cars targeted for freeway entry and the actual number of cars that enter. Basically, this adjustment subtracts the number of HOV entries per minute from the metering rate *(5)*.

The existing ramp metering algorithm is primarily based on estimates of the present and predicted vehicular storage rates of the freeway downstream of the ramp meter *(5)*. The existing ramp metering algorithm has a time lag between operational traffic problem detection and corrective action. For instance, a reaction to existing congestion may result in overly restrictive metering rates. Excessive queue build-up may consequently activate the queue override, which increases the metering rate to keep cars from backing up into the arterial. The resulting increase in freeway congestion may then cause the cycle to repeat *(6)*. Once the freeway starts oscillating between restrictive and high metering rates, it may have trouble escaping the cycle until the congestion naturally dissipates. The current algorithm also depends strongly on loop detector data. Induction loops, located under the freeway pavement about every 0.8 km, sample freeway data, which a central computer receives every 20 seconds. Loop detector data are not always reliable because of noisy signals, transmission problems, construction work, and mechanical failure, so effective control responses dependent solely on reliable input data may not always be provided.

One phase of an extended research effort has recently been completed to develop a predictive ramp metering algorithm to overcome the limitations of Seattle's existing ramp metering algorithm *(6)*. An artificial neural network (ANN) was created to predict freeway volume and occupancy during heavily congested flow. This data prediction will provide an input to a fuzzy logic ramp metering algorithm that also was developed. The desired result would be a ramp metering rate that is based on both current and predicted traffic flow. Ideally, the new algorithm may help prevent bottlenecks rather than simply react to them. By considering the freeway as a control system instead of one freeway section at a time, the fuzzy logic algorithm should avoid producing an oscillatory ramp metering rate and should achieve equilibrium more quickly and smoothly.

There were two stages to this research project *(6)*: the development of a neural network traffic data predictor and the development of a fuzzy logic ramp metering algorithm. This project focused primarily on the ANN traffic data predictors, but it also laid the groundwork for the fuzzy logic ramp metering concept and algorithm. Because the fuzzy logic algorithm testing and fine-tuning requires having a calibrated freeway traffic simulation model capable of closed-loop control which is unavailable even at this time, fuzzy logic algorithm testing will require further research and development. A proposal by WSDOT-Seattle to FHWA to provide such features to the FREESIM freeway simulation model are being considered at this time.

3.2 ARTIFICIAL NEURAL NETWORK

A multi-layer perceptron type ANN was trained using back propagation techniques to minimize the mean squared error of the prediction. Study time intervals of data included 20 seconds, 1.0 minutes, and 2.0 minutes at two measurement sites using historical data. Training data were collected to tune the algorithm and other data were collected at the sites to test the generalization capabilities of the trained ANN to predict forthcoming traffic results.

Although the ANNs were trained and tested on historical data, they can be implemented for real-time prediction because the inputs are past values of volume and occupancy. For on-line implementation, a neural network would need to be trained from data for each prediction site. The training algorithm and neural network architecture should remain similar for different sites. The ANN parameters should remain similar as long as the data characteristics are similar to the original sites. If not, the neural network should be trained on-line to allow for seasonal variations. A watchdog should constantly monitor the accuracy of the prediction to indicate whether retraining is necessary.

While the WSDOT researchers were cautiously optimistic about the ANN's ability to predict traffic volume and occupancy 1.0 minutes into the future, they noted serious difficulties in predicting even 2.0 minutes into the future; to quote, "predictions over 1 minute were unreliable. Because of the somewhat chaotic nature of freeway traffic data, longer term prediction with ANNs is a much more difficult problem" *(6)*.

3.3 FUZZY LOGIC RAMP METERING

Fuzzy logic allows the use of qualitive knowledge. Rather than forcing a specific yes or no, on or off response, fuzzy logic utilizes imprecise information such as maybe, occasionally, and probably. Fuzzy Logic Control (FLC) is becoming common in even household control systems. FLC involves four main steps: fuzzification, rule evaluation, implication, and defuzzification. The final FLC ramp metering model, as yet to be implemented and tested, would be tuned by trial and error usage by WSDOT freeway operation personnel.

FLC is thought to be well-suited for ramp metering applications for several reasons. It requires a mathematical model of the system, and it can utilize imprecise or incomplete information. These traits are important, given that freeway operations are difficult to accurately model and that loop detector data are susceptible to error. The fuzzy logic rules incorporate human expertise, considering all factors simultaneously rather than making a series of local adjustments. For easy algorithm modification and code simplicity, adjustable parameters define membership classes for control. A weight for each rule allows that rule to be emphasized or eliminated for tuning purposes. The fuzzy logic algorithm was also designed to overcome the disadvantages of Seattle's current ramp metering algorithm. The parallel rules of the controller promote robustness to faulty loop detector data. Fuzzy logic control can provide smooth transitions rather than threshold activiations, as well as prevent queue formation through the use of qualitative queue inputs. Rules based on the premise of low downstream speed and high downstream occupancy provide a better indicator of bottlenecks than does the downstream storage rate algorithm now used by WSDOT.

Fuzzification

The fuzzification process translates each precise input into a set of fuzzy variables defined by membership classes, also called membership functions. A membership function describes, on a scale of 0 to 1, the degree to which an input belongs to that set. Membership functions can be discrete or continuous, and triangles or bell-shaped curves commonly define them. Figure 3.1 shows membership classes appropriate for the storage rate, that is, the number of vehicles entering a section minus the number of vehicles exiting a section during the past minute. Each fuzzy variable represents a class as follows:

NB - negative big
NS - negative small
ZE - zero
PS - positive small
PB - positive big

If the numerical, or crisp, storage rate input is 12 vehicles/minute, then the fuzzy PS degree is 0.6, and the PB degree is 0.2, with the remaining classes zero. If the crisp input is less than -20 vehicles/minute, the NB membership function is 1, and if the crisp input is greater than 20 vehicles/minute, the PB membership function is 1. If more membership classes are added, such as negative medium, the triangular bases can be narrowed. With fewer membership classes, more overlap is needed. The best percentage of overlap between classes depends on the specific application. Researchers recommend between 25 percent overlap (Kosko, 1992) and 75 percent overlap (Lin and Lee, 1993; Gupta, 1991).

Rules

The rules, sometimes called the knowledge base, are the heart of a fuzzy logic controller. Rules are based on expert opinions, operator experience, and system knowledge. For the fuzzy logic ramp metering algorithm, the existing ramp metering algorithm provides a starting point for rule development. Rule evaluation, based on fuzzy set theory, uses fuzzy operators to perform logical

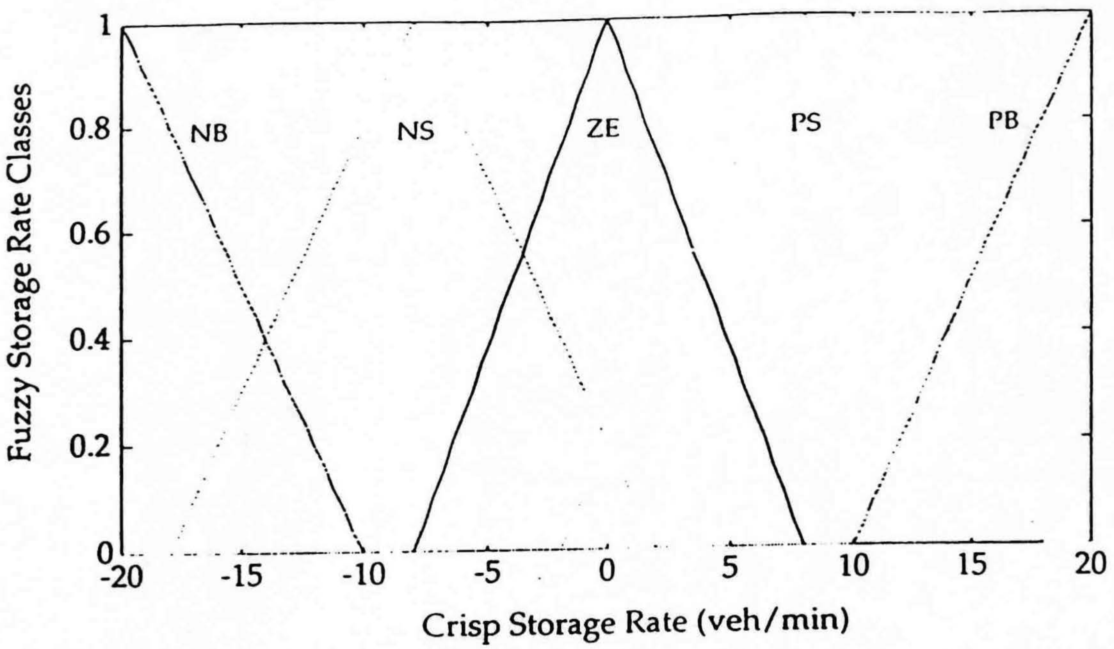

Figure 3.1 Storage rate membership classes.

operations such as the complement, intersection, and union of sets. Complementation corresponds to one minus the membership degree in fuzzy set theory. An AND operation, analogous to the intersection of sets, takes the minimum of given membership degrees. An OR operation, analogous to the union of sets, takes the maximum of the given set of membership degrees.

A simplified example for a fuzzy logic ramp metering algorithm demonstrates rule evaluation. Suppose a knowledge base includes the following rules:

Rule 1:	If downstream storage rate is positive big and downstream occupancy is big, then metering rate is small.
Rule 2:	If queue override occupancy is big and queue duration is big, then metering rate is small.
Rule 3:	If upstream occupancy is small, then metering rate is big.
Rule 4:	If predicted occupancy is small, then metering rate is big.

These rules can be rewritten more compactly using the variables given in Table 3.1. The variables in parenthesis below represent the qualifying conditions, and the number in brackets is a hypothetical membership degree. This conditional pair is followed by the output metering rate MR to the degree shown in brackets:

Rule 1: If (SR_PB [.5], DO_B [.7]) then MR_S [.5]
Rule 2: If (QO_B [.6], QD_B [.2]) then MR_S [.2]
Rule 3: If (UO_S [.3]) then MR_B [.3]
Rule 4: If (PO_S [.2]) then MR_B [.2]

Table 3.1 Variable Descriptions for FLC Example

Variable	Description
SR	downstream storage rate
DO	downstream occupancy
QO	ramp queue occupancy
QD	ramp queue duration
UO	upstream occupancy
PO	predicted mainline occupancy
MR	ramp metering rate

Two methods for further reducing rules with similar outputs are the *maximum and additive* methods. The *maximum* method of rule deduction takes the maximum MR_B degree of rules 1 and 2, since this corresponds to a union of the two output sets. Using the maximum method, rules 1 and 2 further reduce to

Composite Rule 1 and 2:

If (SR_PB [.5], DO_B [.7]) OR (QO_B [.6], QD_B [.2])
then MR_S [.5]

Similarly, the maximum method combines Rules 3 and 4 to

Composite Rule 3 and 4:

If (UO_S [.3]) OR (PO_S [.2]), then MR_B [.3]

The additive method adds the two output degrees together. With this method, rules 1 and 2 produce MR_S [.7], and rules 3 and 4 produce MR_B [.5]. These two rule deduction methods produce different results, and the one that is most appropriate depends on the application. At this point, each output variable class is implicated to a degree.

Implication

Implication expresses the area that an output variable class activates for use in the defuzzification calculation. Two common implication mechanisms are *correlation-minimum* encoding and *correlation-product* encoding. The *correlation-minimum* method uses the min operator,

$$\min(w, f(x) \qquad (3.1)$$

Which simply cuts off the class, $f(x)$, at the output degree, w. The *correlation-product* encoding method scales the output area by the output degree:

$$w * f(x) \qquad (3.2)$$

Figure 3.2 demonstrates these two implication methods graphically for an output degree of 0.5. The shaded area represents the implicated area. These two methods may produce different results, but correlation-product implication can make defuzzification easier.

Defuzzification

The defuzzification process produces a crisp output given a fuzzy output variable set. The commonly used centroid method finds the crisp output by dividing the sum of the implicated areas into two equal areas:

$$\frac{\int xf(x)\,dx}{\int f(x)\,dx} \qquad (3.3)$$

Figure 3.3 illustrates the centroid method for the previous ramp metering example. Using the correlation-minimum inteference mechanism, the sum of the MR_S [.5] and MR_B [.3] implicated areas produces a crisp metering rate of 6 vehicles/minute.

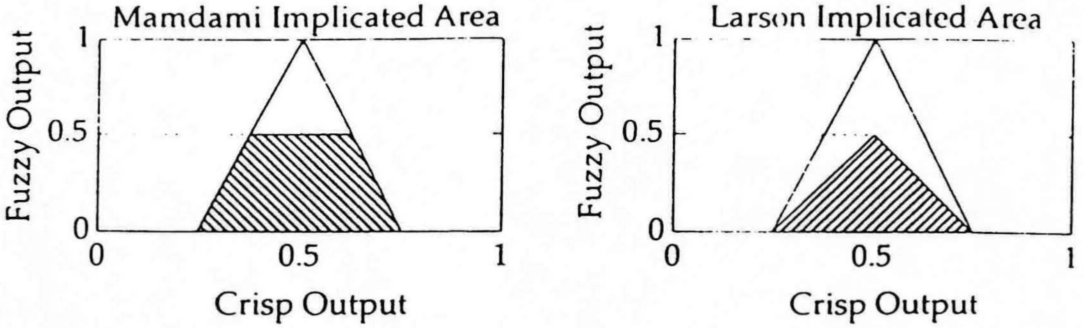

Figure 3.2 Correlation-minimum (left) and Correlation-product (right) implication methods.

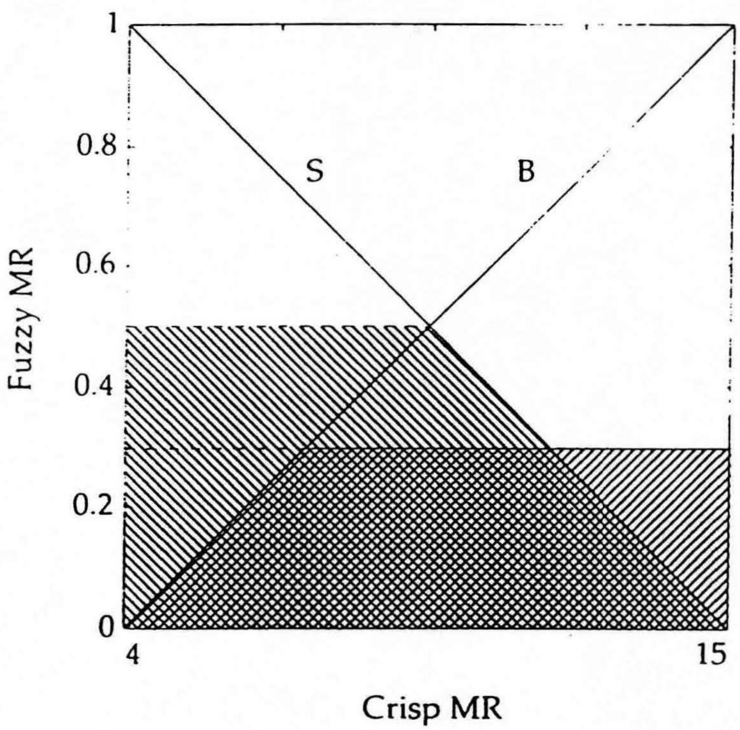

Figure 3.3 Centroid defuzzification.

In practice, a discrete fuzzy centroid is easier to calculate than the above continuous centroid equation. For the case in which correlation-product inference is used, the following discrete centroid equation is equivalent to the continuous centroid equation (see Kosko, 1992 for proof):

$$\frac{\sum_{i=1}^{N} w_i c_i I_i}{\sum_{i=1}^{N} w_i I_i} \tag{3.4}$$

where c_i is the centroid and I_i is the area of the output class for the ith rule.

3.4 FUZZY LOGIC RAMP METERING ALGORITHM

Flexibility is a key issue in the design of a fuzzy logic ramp metering algorithm. For reasons that will be discussed, WSDOT is contracting out the work to incorporate the fuzzy logic algorithm into a freeway model. For this reason, additional source code modifications would be time consuming, so the initial algorithm had to be designed to be flexible. Because the purpose of model testing is to tune the algorithm by trial and error, the class definitions, and rules have to be designed for easy modification.

It is desirable to have a minimum number of inputs and rules. The initial fuzzy ramp metering algorithm contains a reasonable number of inputs and rules that may be useful for control, and the testing and tuning process with a model will determine which of these inputs and rules are useful. The inputs and rules that are unnecessary may be eliminated without making extensive source code modifications.

Table 3.2 describes the variables used as inputs to the fuzzy logic controller, and Figure 3.4 shows the location of each input variable. All input variables are based on a 20-second sampling period, unless stated otherwise in the table. Mainline variables are based on loop detectors across all lanes. For example, volume is in vehicles per 20 seconds accumulated across all lanes. The predicted occupancy is a neural network output.

Table 3.2 Description of Algorithm Variables

Variable	Description
VO	mainline volume just before ramp merge
OC	mainline occupancy just before ramp merge
DO	downstream occupancy of nearest bottleneck prone section
UO	upstream occupancy for adjacent station
PO	1-minute prediction of mainline occupancy just before ramp merge
SP	speed for mainline just before ramp
DS	downstream speed of nearest bottleneck prone section
SR	downstream storage rate of nearest bottleneck prone section
QO	ramp queue occupancy over the past sample
QD	ramp queue occupancy averaged over the past 6 samples
AQO	advanced queue detector occupancy over past sample
AQD	advanced queue detector occupancy average over past 3 samples
MR	metering rate (the control action)

Sample time: 20 seconds.

Table 3.3 Fuzzy Classes

I	Class	Description
1	NB	negative big
2	NS	negative small
3	ZE	zero
4	PS	positive small
5	PB	positive big

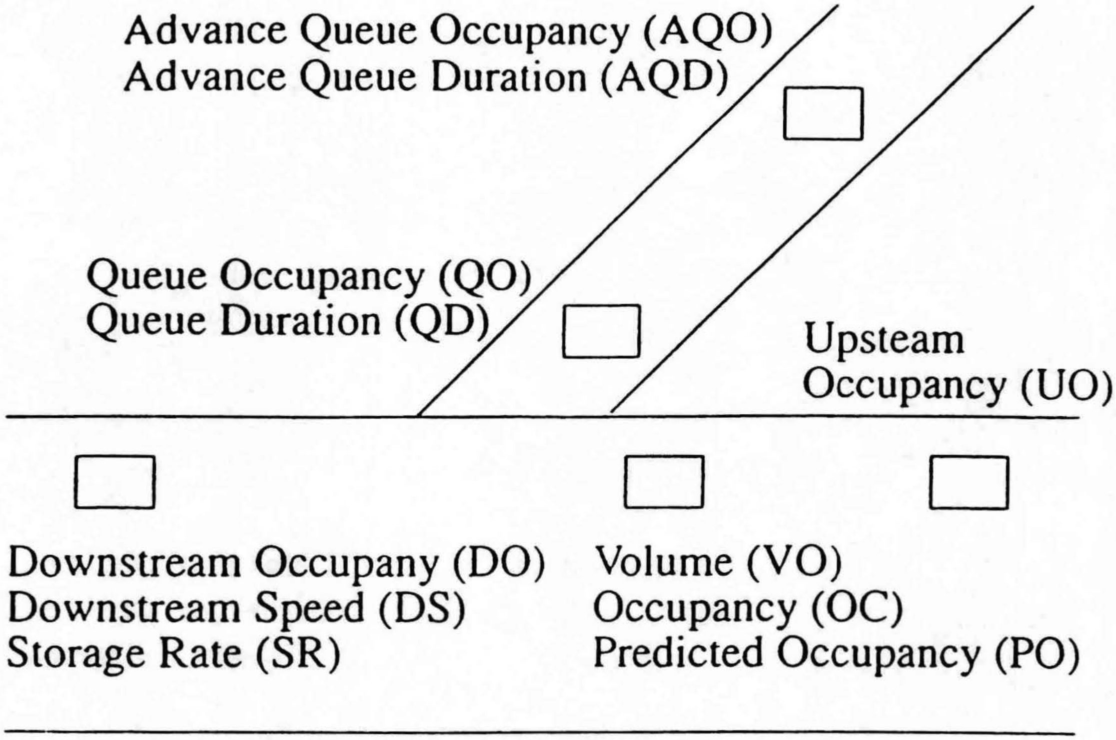

Figure 3.4 Location of algorithm variables.

Each variable has several parameters that add flexibility to the class definitions. The first two scaling parameters set the low limit (LL) and high limit (HL) for the dynamic control range of each variable. The following scaling equation normalizes the crisp variables from the (LL,HL) range to the (0,1) range:

$$\textit{scaled crisp variable} = \frac{\textit{crisp variable}}{HL-LL} - \frac{LL}{HL-LL} \tag{3.5}$$

The scaling simplifies the code by allowing all variables to use the same fuzzification equations, as well as allowing easy class modification. By increasing LL and decreasing HL, the sensitivity to a particular input can be increased, which causes the rules with that premise to fire to a greater degree.

The fuzzification process translates each controller input variable into 5 classes (Table 3.3). In addition to limit parameters, each variable class has a centroid (C_i) and width (B_i) parameter to define the ith class. The NS, ZE, and PS classes are defined by an isosceles triangle with a base of $2\beta_i$ and a height of 1. The triangle is centered at C_i and has slopes of $\pm \frac{1}{\beta_i}$. The resulting fuzzy are defined by the scaled crisp variable, x. For NS, ZE, and PS,

$$f(x) = \begin{cases} \frac{1}{\beta_i}(x-C_i+\beta_i) & \textit{for } C_i-\beta_i < x < C_i \\ -\frac{1}{\beta_i}(x-C_i-\beta_i) & \textit{for } C_i < x < C_i + \beta_i \end{cases} \tag{3.6}$$

Table 3.4 Parameter Input Card

Fuzzification Parameters

		LL	HL	C_NS	C_ZE	C_PS	B_NB	B_NS	B_ZE	B_PS	B_PB
V	VO	150	185	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
a	OC	8	18	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
r	DO	8	18	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
i	UO	8	18	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
a	PO	8	18	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
b	SP	45	65	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
l	DS	45	65	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
e	SR	-15	15	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
s	QO	10	60	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
	QD	10	60	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
	AQO	5	10	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
	AQD	5	10	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25
	MR	2	5	0.3	0.5	0.7	0.25	0.25	0.2	0.25	0.25

Rule Weights

Rule #	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Weight	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Rule #	15	16	17	18	19	20	21	22	23	24	25	26	27
Weight	1	1	1	1	1	1	1	1	1	1	1	1	1

As in Figure 3.3, a right triangle defines NB and PB. For NB, The peak is at O, so C_i is β_i. The class is 1 if x is less than o. For NB,

$$f_i(x)=\begin{cases} 1 \; for < 0 \\ -\frac{1}{\beta_i}(x-\beta_i) \; for \; 0 < x < \beta_i \end{cases} \tag{3.7}$$

For PB, the peak is at 1, so C_i is $1 - \frac{\beta_i}{3}$. The class is 1 if x is greater than 1. For PB,

$$f_i(x)=\begin{cases} \frac{1}{\beta_i}(x-1+\beta_i) \; for \; 1-\beta_i < x < 1 \\ 1 \; for \; x > 1 \end{cases} \tag{3.8}$$

An example of parameter values to define the set of fuzzy classes for SR (shown in Figure 3.1) are LL=-20, HL=20, C=[0.083, 0.3, 0.5, 0.7, 0.916], and β = [0.25, 0.25, 0.2, 0.25, 0.25]. Table 3.4 shows a sample input card containing the parameters that are user specified. Default values are shown. One input card is required for each on-ramp that uses fuzzy control metering. Notice that C_NB and C_PB are not user specified, as they can be calculated from B_NB and B_PB.

Like the class definitions for each variable, the rules need flexibility. Hence, each rule is assigned a weight, which can be set to zero to eliminate that rule (See Table 3.4). If a rule is more important than other rules, it can have a greater weight. For example, the advance queue override rules 7c and 7d in Table 3.4 may have a weight of 2, while most other rules have a weight of 1. To avoid future source code modifications, the initial ramp metering algorithm contains a large number of rules which that will be tested in simulations (Table 3.5). By testing this algorithm on a freeway model, this large rule base can be pared down to as few rules as possible. For each rule in the table, the intersection of the premises involves using the minimum of the premise degrees as the output degree. To combine rules that produce the same output class, the additive method is used rather than the *maximum* method. The *additive* method is expected to be less sensitive to faulty loop detector data. While the *maximum* method may "choose" the faulty value because it is the most extreme, the *additive* method will average together each rule contribution. Another factor to consider is that this rule base does not individually consider all of the possible 5^{13} input combinations. Instead, the rule base is completed by specifying a metering rate given any value of occupancy or upstream occupancy. By averaging rule outcomes together, the control action could be smoother.

Related rules are grouped together. Rules 1a through 1e address the concern for the completeness of the rule base. Because the entire range of occupancy inputs is considered, at least one of these five rules should fire. If OC is not available, the predicted occupancy rules 2a and 2b and the upstream occupancy rules 3a through 3e can produce a similar output. Volume is not used as a premise for rule 1 because occupancy is a more reliable indicator of congestion. However, volume is used in the speed calculation, which is also a reliable indicator of congestion. Rules 4a through 4d adjust the metering rate on the basis of mainline speed.

Rules 5 and 6a through 6d are devoted to preventing or delaying downstream bottleneck formation. Rule 5 emulates Seattle's existing bottleneck algorithm, in which high storage rate and high occupancy downstream indicate a bottleneck. However, plots of storage rate show that is not an accurate indicator of congestion (Nihan 95). Storage rate fluctuates around zero whether in light traffic or heavy congestion. This traffic behavior agrees with intuition. Consider that storage rate is the number of vehicles being added to a freeway section during a sampling period. The number of vehicles that can fit into a freeway section has a limit, so the average storage rate over a long time is zero. Even in stop and go traffic, vehicles must still exit the bottleneck section. Because vehicles tend to travel in platoons, the storage rate oscillates rapidly between positive and negative. The fuzzy algorithm uses downstream speed and occupancy inputs, which are superior indicators of mainline congestion (Iwasaki 1991; Masher, Ross, Wong, Tuan, Zeidler and Petracek, 1975)." Rules 6a through 6e use the premise that high downstream occupancy and low downstream speed indicate bottleneck conditions. For an improved bottleneck indicator, rules 6a through 6e can be used in place of rule 5.

Table 3.5 Rule Base for Fuzzy Ramp Metering Algorithm

Rule	Premise	MR outcome
1a	OC_PB	NB
1b	OC_PS	NS
1c	OC_ZE	ZE
1d	OC_NS	PS
1e	OC_NB	PB
2a	PO_PB	NB
2b	PO_NB	PB
3a	UO_PB	NB
3b	UO_PS	NS
3c	UO_ZE	ZE
3d	UO_NS	PS
3e	UO_NB	PB
4a	SP_NB, OC_PB	NB
4b	SP_NS	NS
4c	SP_PS	PS
4d	SP_PB, OC_NB	PB
5	SR_PB, DO_PB	NB
6a	DS_NB, DO_PB	NB
6b	DS_NS, DO_PS	NS
6c	DS_ZE, DO_ZE	ZE
6d	DS_PS, DO_NS	PS
6e	DS_PB, DO_NB	PB
7a	QO_PB	PS
7b	QD_PB	PB
7c	AQO_PB	PB
7d	AQD_PB	PB

Rules 7a through 7d in Table 3.5 address ramp queue occupancy and duration. Seattle's current algorithm is susceptible to cycling between restrictive and high metering rates during peak hours. Rules 7a through 7d provide smooth transitions rather than threshold activations. When the queue occupancy is high, the metering rate increases. A high queue duration over the past six samples indicated that the queue is building up over time. Unlike the threshold activation of Seattle's current algorithm, the queue duration input provides qualitative information regarding the queue formation. With this information, rules 7a through 7d can help prevent queue formation and avoid an oscillatory metering rate. The advance queue override detector for rules 7c through 7d is located closer to the arterial than the queue override detector for rules 7a and 7b, so rules 7c and 7d should be weighted more heavily to prevent vehicles from backing up into the arterial.

For calculation simplicity, the implication method used is *correlation-product* encoding rather than *correlation-minimum* encoding. Correlation-product inference allows use of the discrete centroid equation,

$$\frac{\sum_{i=1}^{N} w_i c_i I_i}{\sum_{i=1}^{N} w_i I_i} \tag{3.9}$$

where c_i is the centroid and I_i is the area of the output class for the ith rule. Given that this FLC has a large number of inputs and rules and requires flexibility, the discrete centroid equation is used for code simplicity. For MR classes of NS, ZE, and PS, the class area I_i equals β_i. For MR classes of NB and PB, the class area I_i equals $\beta_i/2$. Once the centroid of the crisp MR has been found, this control action must be rescaled back to its dynamic range using the scaling equation. Like Seattle's bottleneck algorithm, the MR should then be adjusted to account for the number of HOVs during the previous sampling interval. The maximum MR possible is 900 vehicles/lane/hour to allow a 4-second cycle for each car, when one car is released at a time. The minimum MR is 240 vehicles/lane/hour, for a maximum vehicle service delay of 15 seconds. Drivers are apt to run a ramp metering light if it is red for more than 15 seconds. Thus, reasonable limits for headway (defined as the cycle length of the light sequence) are 4 and 15 seconds. The metering rate is equal to the sampling period divided by the headway, for limits of LL=2 and HL=5 vehicles/sample. The fuzzy logic controller code written in C is provided in Appendix C of the Seattle report *(4)*.

4. ALINEA: LOCAL FEEDBACK CONTROL LAW FOR RAMP METERING

4.1 PROCESS CONTROL

Consider a process under control (e.g., home heating), and a selected (or desired) output (e.g., inside room temperature), shown schematically in Figure 4.1 *(7)*. The process is affected by some process inputs. Process inputs that can be manipulated are called controllable inputs, or simply inputs (e.g. heating valves); whereas, process inputs that affect outputs, but cannot be manipulated are called disturbances (e.g. outside temperatures). Disturbances may be predictable or nonpredictable, measurable or nonmeasurable, etc. The control problem is to appropriately select the controllable inputs so as to achieve (despite the impact of disturbances) a desired process output value called the "set value" (e.g., a desired room temperature of 70° F).

Assume that a mathematical model of the process is available, and furthermore that all essential disturbances are time variant but measurable. Then, at any instant of time, inputs can be calculated on the basis of the process model using disturbance measurements to achieve a given set value (Figure 4.1a). This feedforward control procedure is broadly known as disturbance compensation. Because of inevitable inaccuracies of mathematical models and occurrence of other unexpected, nonmeasurable disturbances, disturbance compensation is known to be a particularly sensitive control structure. In the home heating example, disturbance compensation would correspond to measuring the outer temperature and controlling the heating valves to achieve a constant inner temperature of, say 70°F.

Consider a traffic flow process around an on-ramp as shown in Figure 4.2. Assuming the absence of congestion, the downstream traffic volume q_{out} may be declared as the process output with a set value (target) $\hat{q}$ (e.g., equal to capacity), whereas the on-ramp volume r is identified as the controllable input and the upstream traffic volume q_{in} as a measurable disturbance, but not an input as traditionally viewed in ramp metering systems. In order to keep q_{out} near $\hat{q}$, an intuitive way to do it is to calculate $r = \hat{q} - q_{in}$ using current measurement of the disturbance q_{in} . Obviously, this feedforward procedure, which is essentially applied by many known local methods for ramp metering, corresponds to the disturbance compensation of Figure 4.1a. The quality of results depends on the accuracy of the applied process model, but complicated models lead to complicated control algorithms. However, even for highly complex strategies, sensitivity with respect to existing inaccuracies and unexpected disturbances remains a structural drawback. Moreover, adjustment of threshold parameters during implementation becomes a difficult task for complicated strategies.

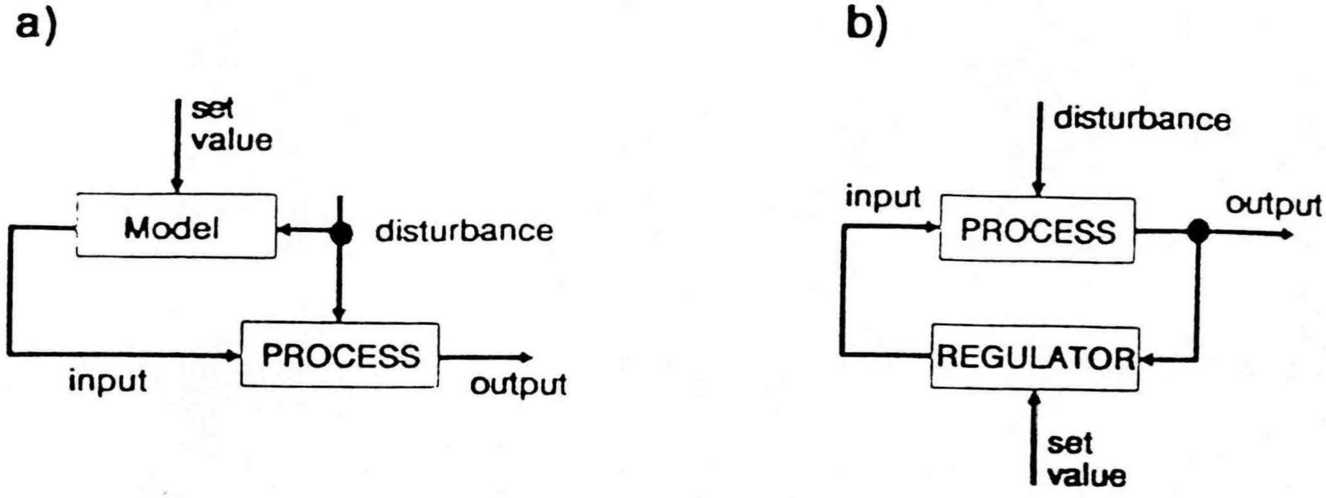

Figure 4.1 (a)Disturbance compensation and (b) feedback control.

A much more elegant, robust, simple, and efficient way of solving the control problem is to introduce a feedback structure (Figure 4.1b). The measurable output is fed back to the system, and the controllable input is permanently modified by an appropriate regulator to keep the output near its set value despite the influence of time-variant disturbances. Design of the regulator may be performed by use of well-known automatic control methods. Because of its feedback structure, a control system of this kind is much more precise and much less sensitive with respect to model inaccuracies and unexpected disturbances as compared with disturbance compensation. In the heating example, feedback control corresponds to measuring the inner temperature of the house and modifying the heating valves accordingly to achieve the desired set value despite the variations of the outside temperature.

The feedback methodology can now be transmitted to the problem of ramp metering shown in Figure 4.2. Because the same feedback law is to be applied both for congested and for free flowing traffic, it is preferable to consider occupancy o_{out} as an output variable instead of q_{out}. This is because traffic volume may have the same values both for light and congested traffic because of the characteristic parabolic form of the fundamental diagram. The corresponding set value $\hat{o}$ for traffic occupancy may be easily found on the basis of the fundamental diagram at the output line; alternatively, a desired downstream occupancy value may be provided directly *(8)*.

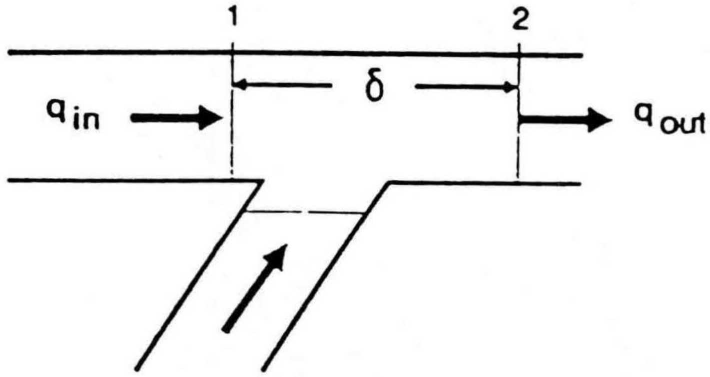

Figure 4.2 A traffic flow process.

An additional advantage of choosing o_{out} rather than q_{out} as an output variable arises from the fact that the critical occupancy o_{cr} seems to be less sensitive with respect to weather conditions and other operational influences compared with the capacity q_{cap} of a freeway stretch. This statement is supported by data material provided by Keen et al. *(9)*. As a consequence, considering the set value $\hat{o} = o_{cr}$ is a more robust way of achieving capacity flow than considering $q = q_{cap}$ because variations of q_{cap} caused by environmental or other conditions are stronger compared with variations of o_{cr}.

Thus, the next step is to derive a feedback control law $r = R(o, o_{out})$ according to Figure 4.1b to keep o_{out} near $\hat{o}$. Derivation of this feedback law is the subject of the next section.

4.2 DERIVATION OF THE FEEDBACK LAW

Modeling

Consider the traffic flow process shown in Figure 4.2. Site 1 is assumed to be situated just upstream of the on-ramp. Site 2 is situated downstream of the on-ramp, at a distance δ from Site 1. Assume that the on-ramp volume r is updated every T time units, where $T = 20$ sec... 1 min, or more. The conservation equation for the freeway stretch between Sites 1 and 2 (Figure 4.2) is

$$p(t) = [q_{in}(t) + r(t) - q_{out}(t)]/\delta \tag{4.1}$$

where the traffic density p (veh/km) is defined as the number of cars included in the stretch, divided by length δ, t being the time argument. Because traffic density is not readily measurable, it is convenient to replace $p(t)$ in Equation 4.1 by the occupancy $o_{out}(t)$ using the approximate relationship $p = \alpha o_{out,}$ where $\alpha = \mu/100\lambda$). μ being the number of lanes of the mainstream and λ being the mean effective vehicle length in kilometers.

Assuming $q_{out}(t)$ is given as a nonlinear function of $o_{out}(t)$ (fundamental diagram)

$$q_{out}(t) = Q[o_{out}(t)] \tag{4.2}$$

and substituting Equation 4.2 into Equation 4.1, a nonlinear first-order dynamic system model is obtained. This model may be linearized around a nominal steady state ($\bar{o}_{out}$, $\bar{q}_{in}$, $\bar{r}$) such that

$$\bar{o}_{out} = \hat{o};\ \bar{q}_{out} = Q(\bar{o}_{out});\ and\ \bar{r} = \bar{q}_{in} - \bar{q}_{out} \tag{4.3}$$

With the notation $\Delta o_{out}(t) = o_{out}(t) - \bar{o}$ used analogously for all variables, the linearization yields

$$\Delta o_{out}(t) = [\Delta q_{in}(t) + \Delta r(t) - \hat{Q}'_{out}(t)]/(\delta\,\alpha) \tag{4.4}$$

where $\hat{Q}' = dQ\,(\hat{o})/do_{out}$, i.e., $\hat{Q}'$ is the slope of the tangent of the fundamental diagram at $\hat{o}$ and hence its value is proportional to the speed of the corresponding kinematic traffic wave *(8)*. Clearly, $\hat{Q}'$ is positive on the left-hand side of the fundamental diagram.

Because the control input r is updated every T time units, time discretization of Equation 4.4 with sample time interval T yields

$$\Delta o_{out}(k+1) = \beta\Delta o_{out}(k) + [(1-\beta)/\hat{Q}'] \times [\Delta q_{in}(k) - \Delta r(k)] \tag{4.5}$$

where $k = 0,1,2...$ is the sample time index. Thus $o_{out}(k)$ is the occupancy at time kT, and $\Delta q_{in}(k)$ and $\Delta r(k)$ are assumed to be constant during the time interval $[(k-1)T, kT]$. The constant parameter β results from the discretization procedure to be $\beta = \exp(-\hat{Q}'\alpha T/\delta)$.

The parameter β may be neglected if the ratio δ/T is sufficiently small. This will be the case if traffic volume entering the freeway reaches Site 2 during the time interval T. In this case, the effect of entering traffic $r(k)$ will be visible at Site 2 by the end of the corresponding time interval. Setting $\beta = 0$ in Equation 4.5 yields

$$\Delta o_{out}(k+1) = [\Delta q_{in}(k) + \Delta r(k)]/\hat{Q}' \tag{4.6}$$

In the next section, a feedback law is developed under the assumption $\beta = 0$. An extension applying to the case of higher ratios δ/T will be derived later. Finally, a complementary disturbance compensation mechanism will be presented.

Regulator Design

An appropriate feedback law for the process of Equation 4.6 is given by the following integral regulator

$$r(k) = r(k-1) - K_R[\hat{o} - o_{out}(k)] \tag{4.7}$$

where K_R is a constant positive regulator parameter. After applying this regulator to Equation 4.6, the following z-transfer function is obtained for the closed-loop system:

$$H(z) = o_{out}(z)/\hat{o} = \frac{K_R/\hat{Q}'}{z - 1 + (K_R/\hat{Q}')} \tag{4.8}$$

A time-optimal deadbeat regulator is obtained by choosing $K_R = \hat{Q}'$. Because of the sign of $\hat{Q}'$, the linearization of Equation 4.6 is strictly valid only on the left-hand side of the fundamental diagram. However, even for congested traffic the feedback law of Equation 4.7 leads to traffic occupancy reduction and can thus be applied in the same way in all operational cases.

Extension for Bottleneck Further Downstream

If Site 2 is located at a bottleneck further downstream and if T is chosen accordingly short, the entering traffic may not reach Site 2 at the end of each time interval, in which case β cannot be neglected. Applying the regulator of Equation 4.7 to the original process model of Equation 4. 5, the closed-loop z-transfer function becomes

$$H(z) = o_{out}(z)/\hat{o} = \frac{z(1-\beta)}{z^2 - 2\beta z + \beta} \tag{4.9}$$

with the eigenvalues $\zeta_{1,2} = \beta \pm j[\beta(-\beta]^{1/2}$. Although the closed-loop system remains stable for any $\beta < 1$, the transient behavior may be slow. An amelioration may be achieved by application of the following proportional-plus-integral feedback law

$$r(k) = r(k-1) - K_R[\hat{o} - o_{out}(k)] \; - K_p[o_{out}(k) - o_{out}(k-1)] \tag{4.10}$$

where K_p is a further constant and positive regulator parameter. A time-optimal deadbeat regulator is achieved by the choice $K_R = \hat{Q}'$, $K_p = \beta\ \hat{Q}'/(1 - \beta)$, as can be readily demonstrated. In summary, a further term has been added in Equation 4. 9 as compared with Equation 4. 7 for the particular case of, say $\delta/T > \alpha\ \hat{Q}'$.

Elimination of Constant Disturbances

The regulator of Equation 4.7 is capable of eliminating constant disturbances. The z-transfer function of the feedback law is

$$\frac{r(z)}{-\Delta o_{out}(z)} = \frac{K_R z}{z - 1} \tag{4.11}$$

Considering the process model of Equation 4.6 with $K = 1/\hat{Q}'$, the corresponding closed-loop linearized system is shown in Figure 4.3. Assume a disturbance d as indicated in Figure 4.3. The z-transfer function o_{out}/d is given by

$$\frac{o_{out}(z)}{d(z)} = \frac{K(z-1)}{z(z-1+KK_R)} \tag{4.12}$$

If d is constant, this equation yields in the steady-state (i.e., for z = 1) $O_{out}/d = 0$, i.e., any constant disturbance is eliminated in the steady state by the control system. Because the upstream traffic volume q_{in} acts as the disturbance d in Figure 4.3, this statement holds true for constant (or slowly varying) upstream traffic volumes.

Elimination of Biased On-Ramp Volume Realization

What happens if the implemented on-ramp volume is biased as compared with the on-ramp volume ordered by the regulator? Bias might result from motorists violating the ramp signal at a constant rate. A bias in the on-ramp volume realization corresponds exactly to the constant disturbance d of Figure 4.3. Hence, ramp bias is automatically eliminated by the control system.

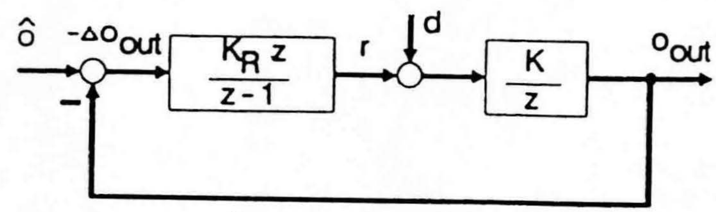

Figure 4.3 Linearized closed-loop system.

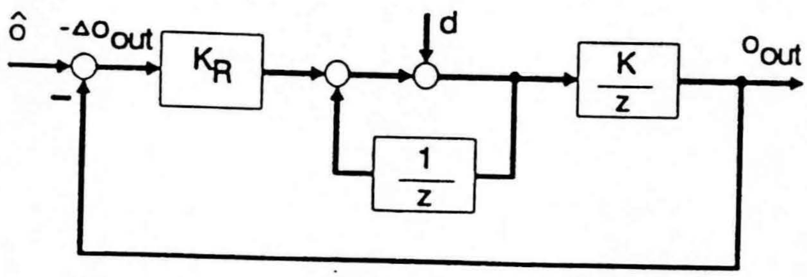

Figure 4.4 Modified closed-loop system.

This statement does not hold true if the implemented on-ramp volume is used for the retarded value $r\,(k-1)$ in the feedback law of Equation 4.7. Figure 4.4 shows the signal flow diagram of the closed-loop system in this case.

The z-transfer function is now given by

$$\frac{o_{out}(z)}{d(z)} = \frac{K}{z - 1 + KK_R} \tag{4.13}$$

If the disturbance d is constant, this equation yields in the steady state $O_{out}/d = 1/K_R$, i.e., a bias $\bar{d}$ leads to a steady-state error (offset) of $\bar{d}/K_R$.

4.3 DISCUSSION OF FEEDBACK CONTROL LAW

The Regulator

The ALINEA *(7)* feedback control law to be applied at the time instants kT, $k = 0,1,2...$, for any sample interval T (e.g., T = 60 sec), is

$$r(k) = r(k-1) + K_R[\hat{o} - o_{out}(k)] \tag{4.14}$$

where K_r is a positive, constant regulator parameter and $o_{out}(k)$ current downstream (not upstream) occupancy measurement. The feedback law of Equation 4.10, which is much simpler than any other local metering strategy, was given the name asservissement Linéaire d' entrée autoroutière (ALINEA) by the cited researchers *(7)*.

Heuristic Interpretation

Equation 4.10 suggests a fairly plausible control behavior. If the measure occupancy o_{out} (k) at time k is found to be lower (higher) than the desired occupancy $\hat{o}$, the second term of the right-hand side of Equation 4.10 becomes positive (negative) and the ordered on-ramp volume r (k) is increased (decreased) as compared to its last value r $(k$ - 1). Clearly, the feedback law of Equation 4.10 acts in the same way both for congested and for light traffic (no switchings are necessary).

Note that some occupancy control strategies *(10)* react to excessive occupancies only after congestion is created and has reached an upstream measurement location, whereas ALINEA reacts smoothly even to slight differences $\hat{o}$ - o_{out} (k) and may thus prevent congestion in an elegant way.

On the other hand, some demand-capacity strategies react to excessive downstream occupancies only after a threshold value is exceeded. Typically, and in contrast to ALINEA, the reaction of these control strategies to excessive occupancies is rather crude: on-ramp volumes are set equal to their minimal values. In this way, a nonnecessary underload of the freeway may occur. On the contrary, the essential effect of ALINEA is to stabilize traffic flow at a high throughput level and eventually to reduce the risk of a breakdown without underloading the freeway *(7)*.

The Value of K_R

A value of $K_R = 70$ veh/hr was found by the researchers to yield good results in real-life experiments. K_R is the only parameter to be adjusted in the implementation phase because no thresholds or other constraints are included in Equation 4.10. Moreover, from theoretical considerations,

- Results are insensitive for a wide range of K_R values;
- Increasing (decreasing) K_R values lead to stronger (smoother) reactions of the regulator, and regulation times get shorter (longer); and
- For extremely high values of K_R, the regulator may have an oscillatory, unstable behavior.

In view of these statements, real-life calibration of the unique free parameter K_R if at all necessary - is particularly easy.

Measurements

ALINEA requires only one detector station that measures occupancy o_{out} downstream of the merge area. This is equal or less than the measurement requirements of other local control strategies. The measurement location should be such that any congestion originating from excessive downstream on-ramp volumes be visible in the measurements. A distance of 40m downstream of the on-ramp nose was found to be adequate at the Boulevard Périphérique in Paris *(7)*. The strategy was also found to work adequately with a distance of 400 m in Amsterdam *(11)*.

Disturbance Reduction

If upstream traffic volume q_{in} is constant, then the feedback law of Equation 4.10 is easily demonstrated to lead to $o_{out} = \hat{o}$ in the steady state. In other words, whatever the value of a constant (and not measured) upstream traffic volume might be, the feedback law leads occupancy to its desired value.

Similarly, if upstream traffic volume q_{in} is perturbed around a constant or slowly varying average, the feedback law of Equation 4.10 keeps downstream occupancy o_{out} close to $\hat{o}$ in the average. On the other hand, rapid oscillations of q_{in} around an average value are only slightly reduced by the control system. Field measurements and testing are needed to fully assess these limitations.

Embedding in a Coordinated Control System

Realization of the feedback law of Equation 4.10 requires a set (target) value $\hat{o}$ be provided by the user. The set value may be changed any time it is needed and hence, the feedback law may be embedded in a simple and natural way into a hierarchical control system with set values of the individual sections being specified (and changed in real time) by a superior coordination level or by an operator (R). Rambo II seeks to provide the coordinated system set point values *(8)*.

Restrictions, Override Tactics

The on-ramp volume values resulting from Equation 4.10 may be limited if some maximum or minimum values are exceeded. Moreover, override tactics (e.g., for preventing interference of the on-ramp queue with surface traffic) may be applied. When either a limitation or override tactic becomes active, the last on-ramp volume $r\ (k - 1)$ required in the feedback law of Equation 4.10 for calculating $r(k)$ should correspond to the actual number of cars that entered the freeway(because of the limitation or override) and not to the calculated but suspended value provided by Equation 4.10 in the last time interval.

Realization of On-Ramp Volumes

ALINEA is compatible with any kind of realization of the required on-ramp volumes (one-by-one, platoon, traffic cycle, etc.). Obviously, the implemented on-ramp volumes should be equal to the on-ramp volumes ordered by the control law of Equation 4.10 on the average unless a limitation becomes active according to the previous section. If this is not true, e.g., if for any reason the implemented on-ramp volumes are biased as compared with the calculated on-ramp volumes, then the control system is capable of eliminating the bias automatically. For this to be true, the value of $r\ (k - 1)$ required in Equation 4.10 should be equal to the on-ramp volume calculated by Equation 4.10 in the last time interval and not equal to the on-ramp volume that actually entered the freeway, unless an override or a constraint has become active, as already mentioned.

Efficacy

A preliminary version of ALINEA and some popular previous control strategies have been implemented and tested on an on-ramp of the Boulevard Périphérique in Paris during an experimentation period of 6 months. Results of this lengthy experimentation period showed a clear superiority of ALINEA in preventing congestion and increasing traffic throughput as compared to other local traffic responsive strategies. Details were reported by Hadj-Salem et. al. *(12,13)*. More recent field results from the Netherlands confirm the superiority of ALINEA as compared with the demand-capacity type of strategies *(11)*.

Extensions

Two possible extensions may be envisaged if the ratio δ/T is relatively high. This condition may hold if the output occupancy o_{out} is measured at a site far downstream of the on-ramp nose or the sample time interval is short (e.g., $T = 10$ sec). As a result of the theoretical development, these extensions are recommended if $\delta/T > \alpha K_R$, where α was defined earlier.

The first extension is to use the feedback law of Equation 4.9 rather than that of Equation 4.10 with the parameter value $K_p = \delta/(t\alpha) - K_R > 0$. Thus, even with this extension K_R remains the only parameter to be calibrated during implementation.

The second extension for the case $\delta/T > \alpha K_R$ is to add to the feedback laws of Equation 4.9 or 4.10 the term $\gamma\,[q_{in}(k) - q_{in}(k-1)]$ if the value $q_{in}(K)$ can be predicted accurately enough by upstream measurements. This second extension corresponds to a disturbance compensation aiming at improving the feedback efficiency. The positive smoothing parameter $\gamma \leq 1$ should be appropriately chosen to avoid excessive oscillations of the on-ramp volumes caused by noise of the q_{in} measurements, a plausible situation that might be encountered.

However, if the output Site 2 is being located far downstream of the on-ramp (at a downstream bottleneck) without any intermediary measurements, there is a risk of congestion being built up (because of incidents or other disturbances) in the interior of the stretch *(12)* without being visible at the output measurement Site 2. In such cases, the feedback (or any other) control system may be of little help for eliminating the congestion if no additional measurement stations are added.

5. A NEURAL NETWORK RAMP METER CONTROLLER

5.1 INTRODUCTION

Model reference control has been a popular approach in control systems for the last several decades. This type of control is supposed to operate satisfactorily in the presence of environmental and parameter uncertainties. Most of this theory, however, is directed toward linear, time-invariant systems. A nonlinear approach based on neural nets is proposed here to control the traffic system described by *(14,15)*. The structure of the nonlinear controller is strongly motivated by that in the linear case. Instead of a linear gain, however, a nonlinear gain is used in the form of a neural net *(14)*. Fig. 5.1 shows the general controller structure *(14)*.

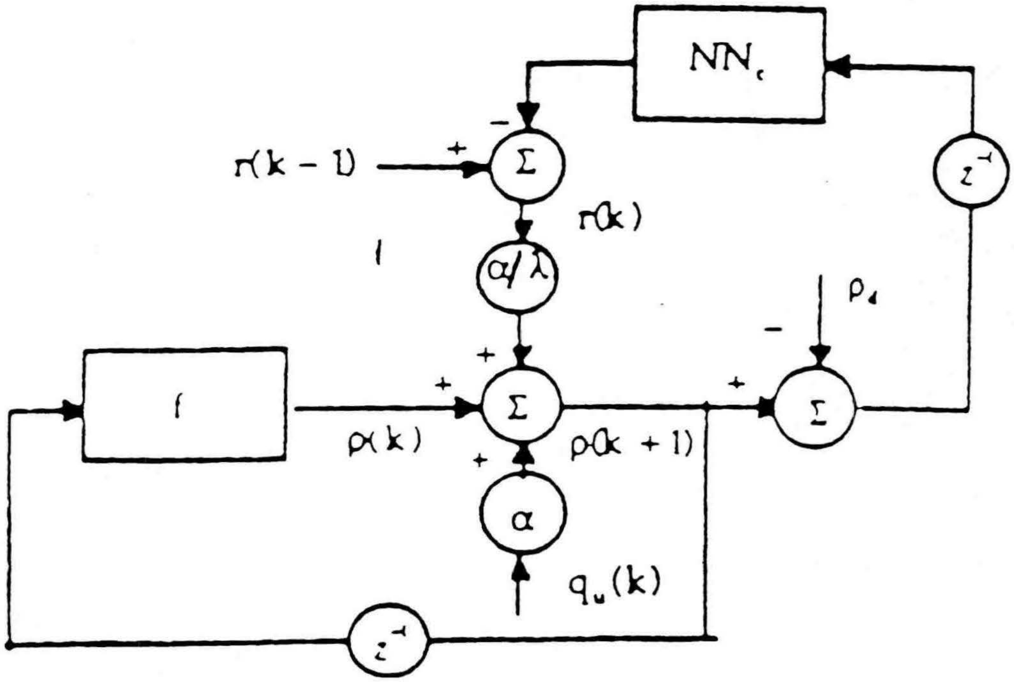

Figure 5.1 The control structure of the neural net controller.

The objective of local ramp metering is to maintain a desired level of service for the freeway system being controlled, such that the freeway system is utilized as fully as possible. To achieve this objective, it is desirable to maintain the local traffic density in a small neighborhood of critical density P_c. Therefore, the reference signal is chosen as:

$$\hat{P}(k) = P_d \tag{5.1}$$

where P_d is the desired traffic density state, which is close to the critical density P_c, and the tracking error is defined as:

$$err(k) = p(k) - P_d \tag{5.2}$$

5.2 CONTROL FUNCTION

Motivated by the PID (proportional-integral-derivative) control concept for linear, time-invariant systems, we prescribe the integral feedback control law *(14)* as:

$$r(k) = r(k-1) - \Psi(err(k)) \tag{5.3}$$

where Ψ is an unknown nonlinear function that belongs to the L^2 space. To find directly such a Ψ that minimizes the tracking error is a difficult task. Instead, a neural net is used to replace Ψ. Such a neural net, generally known as multilayer feed forward neural net, has been proven capable in theory of approximating any function in the L^2 space to an arbitrary degree of accuracy *(15,Hecht-Nielsen, 1990)*.

$$NN_c = \Phi[W_2 \Phi(W_1 u)] \tag{5.4}$$

where u is the input to the neural net. Φ is a nonlinear operator such that $\phi(x) = \tanh(x)$ and ϕ is the diagonal entry of Φ, and W_1, W_2 are connection weight matrices of dimension 1 x m and m x 1, respectively. The weights of the neural net can be derived according to the well known backpropagation algorithm.

5.3 SIMULATION RESULTS

A number of simulations were carried out to test the performance of the proposed neural net controller. In the simulations, the Greenshields flow-density relationship in *(16)* was used, with parameters $v_f = 60$ miles per hour, and $P_j = 120$ vehicles per mile. This results in a maximum flow of 1800 vehicles per hour at $P_c = 60$ vehicles per mile. The reference signal is chosen to be 55 vehicles per mile, the neural net controller was trained for about 10,000 iterations. After its weights stabilized, the trained neural net was used to perform a number of simulations. Because of its robustness and good performance with respect to real data, the ALINEA ramp control algorithm *(7,Papageorgiou et al, 1991)* was used as a bench mark to evaluate the performance of the neural net controller (we linearized equation (4.1) at p = 55 and obtained an optimal gain K = 5.0 according to ALINEA).

In the first simulation, the researchers used an initial density p(0) = 40 vehicles per mile, which is in the uncongested region on the flow density diagram, and $q_u = 1700$ vehicles per hour. Both the neural net controller and ALINEA were able to stabilize the system around P_d after about 80 intervals. But the neural net controller damps faster than ALINEA.

In the second simulation, the initial density was p(0) = 65 vehicles per mile, which is in the congested region on the flow-density diagram. The same inflow volume was applied as in the first simulation. Similar results were obtained as in the first simulation.

In the third simulation, a random inflow q_u was applied to the traffic system, where q_u is assumed to be uniformly distributed over an interval of [1400,1800]. When this disturbance was large, the neural net controller was able to drive the system back close to its nominal state, while ALINEA failed. It is also worthwhile to note that the oscillations in both controls and system responses were significantly reduced in this simulation scenario, which is closer to real world traffic situations than previous ones.

5.4 CONCLUSION

In conclusion, the neural net controller performs as well as ALINEA in most cases. But when the disturbance is large, the neural net controller appears to perform better than ALINEA due to its nonlinear gain and fast damping. As noted from the controls, when the neural net controller tends to overreact, this problem can be remedied by adding a proportional feedback to the control law. Because the magnitude of the disturbance q_u may in general be much larger than that of the control r, the controllability of the freeway system through local ramp metering is often limited, and in some cases coordinated ramp control would be preferred. Further research is underway in University of California, Irvine to extend the local neural net controller to a hierarchical coordinated control structure *(14)*.

6. FHWA'S FRESIM SIMULATION MODEL

6.1 INTRODUCTION

FRESIM is a microscopic simulation model for analyzing ramp metering and other operational improvements in freeway networks. The model was developed for the Federal Highway Administration (FHWA) and is available as both a stand-alone model and as an integrated component of the TRAF simulation family*(17)*. The model, which is an enhanced version of INTRAS model, can simulate more complex freeway geometries and thus provides more realistic representation of traffic behavior than INTRAS.

Due to it's microscopic nature, the FRESIM simulation model can provide the maximum level of geometric and traffic operational features. The level of detail provided by FRESIM can be used for capacity, merging, weaving and ramp metering studies. The following are the geometric and operational simulation capabilities of the FRESIM model, version 5.

Geometric Simulating Capabilities

- Up to five through-lane freeway mainlines, three-lane ramps and three-lane freeway connectors;
- Lane additions and lane drops can also be simulated;
- Geometric variations in radius of curvature, superelevation and grade can also be accommodated in the model;
- Auxiliary lanes to enter or exit the freeway for the lane changing process can also be represented in the model; and
- Freeway blockage incident can be simulated and thereby, work zone can also be represented.

A version of FRESIM has been combined by FHWA with NETSIM to provide an overall freeway corridor simulation model called CORSIM.

Ramp Metering Strategies

- Various ramp metering methods that can be employed by FRESIM in the freeway simulation process are as follows:

Clock Time Metering

Metering headway in seconds is entered for Clock Time Metering, which is the inverse of the metering rate. A count-down clock is assigned to each on-ramp and the metering signal is set to green when the clock returns to zero.

Demand/Capacity Metering

The capacity of the downstream section of the freeway is coded in this type of ramp metering. An evaluation of current capacity is done at certain intervals, based on surveillance detectors and a metering rate is calculated such that the freeway capacity is not violated. The metering rate, thus obtained, is applied as for clock-time metering.

Speed Control Metering

Speed evaluations, based on the detectors placed in the upstream location of the on ramp, are performed to establish a responsive metering rate. For this, three speed thresholds and their corresponding metering rates are specified. At the end of each evaluation period, the prevailing freeway speed is compared to the input speed minimums to determine the proper metering rate.

Gap Acceptance Merge Control Metering

In this method of ramp control, a minimum acceptable gap in tenths of a second is specified. The gaps between the vehicles in the outside lane are estimated using the detectors and the ramp vehicles are released into the freeway whenever an acceptable gap is found.

- Simulation of freeway surveillance system;
- Comprehensive lane changing model, including for incidents and geometric changes like lane drop, etc;
- Six different vehicle types (two types of passenger cars and four types of trucks) and ten different driver types (ranging from timid to aggressive drivers) can be represented in the model; and
- Heavy vehicle movement can be biased or restricted to certain lanes.

Although FRESIM is a detailed microscopic simulation, it has a few limitations in it's application, which are listed below:

- FRESIM can't simulate HOV operations;
- FRESIM doesn't model bus operations; and
- Measures of Effectiveness

The various Measures of Effectiveness (MOEs) that are produced by FRESIM consists of the following:

Input Data Echo Tables

These tables mainly contain the input data, which is for easy reference to check any coding errors.

Intermediate Statistics Output

This output report consists of various measures of effectiveness regarding the operation of the freeway. These are the intermediate link statistics and the vehicle content tables. The statistics include turn movements, delay/vehicle, average speed, number of lane changes, vehicles discharged etc.

Cumulative Statistics Output

This is the cumulative output of the simulation runs and contain measures of effectiveness like vehicle miles, vehicle-minutes, total delay time, volume, density, speed, lane changes etc.

Fuel and Emissions Output

This is an optional output, which contain total fuel consumption, fuel consumption rates, emission rates of HC, CO, NO etc.

6.2 SIMULATION STUDIES

Several simulation studies have been performed in this research using different input schemes (e.g. with and without ramp metering, with and without an incident, with variable truck percentages, volumes, etc.). The basic network model that was simulated consists of a 6-lane freeway with single lane entrance and exit ramps and is shown in Figure 6.1.

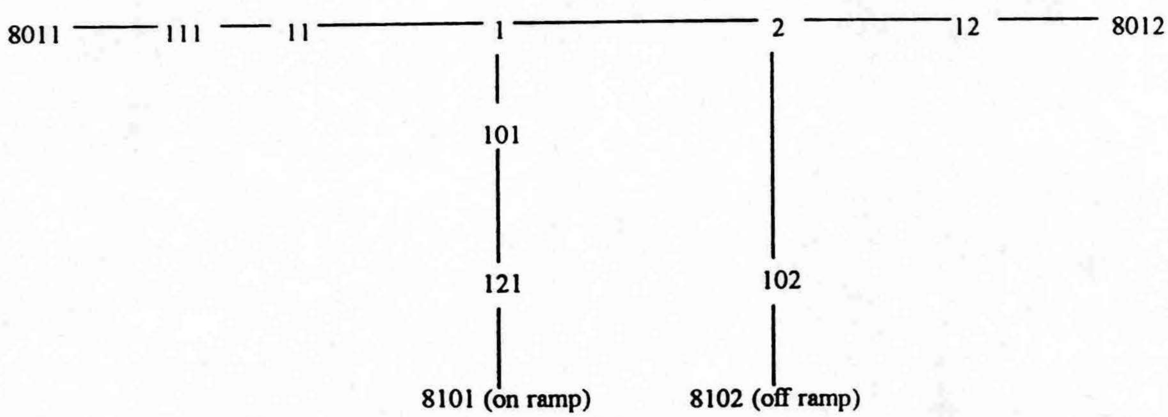

Figure 6.1 Basic ramp metering network.

In Figure 6.1, the freeway links are (11,11), (11, 1), (1,2) and (2,12). Entrance and exit ramps are present at nodes 1 and 2 respectively. The ramp links are (121,101), (101,1) and (2,102). Links consisting of eight thousand series nodes are the entry links. The metering signal is located at node 101 of the on-ramp, which is assumed to be 300 ft away from the node 1 of the freeway. Link length of 2000 ft has been coded for the links (11,1) and (1,2) and a link length of 1000 ft has been input for the ramp link (121,101). The other coded inputs were a freeway speed of 65 mph and a ramp speed of 35 mph. An acceleration lane of length 500 ft and a deceleration lane of length 600 ft have also been coded for the freeway link (1,2). The summary of the simulation studies and results for link 1-2 are as follows.

Case 1 Simulation

In the first case, several simulations were run for various freeway volumes ranging from 1000 vph to 7000 vph without any ramp volume and trucks. The different values of flow, speed and density generated are shown in Table 6.1.

Table 6.1 Case 1 Simulation Results

Total Fwy Volume (vph)	Volume (Veh/ln/hr	Density (Veh/ln-mile)	Speed (Mph)
1000	333	5.2	64.06
2000	666	10.5	63.67
3000	1001	15.9	62.77
4000	1333	21.6	61.64
5000	1667	27.6	60.34
6000	1981	34.2	57.86
7000	2121	38.2	55.5

The Flow vs Speed curve for the freeway link 1-2 was plotted, which is shown in Figure 6.2. The plot shows the relationship between flow and speed for the small speed range noted in Table 6.1.

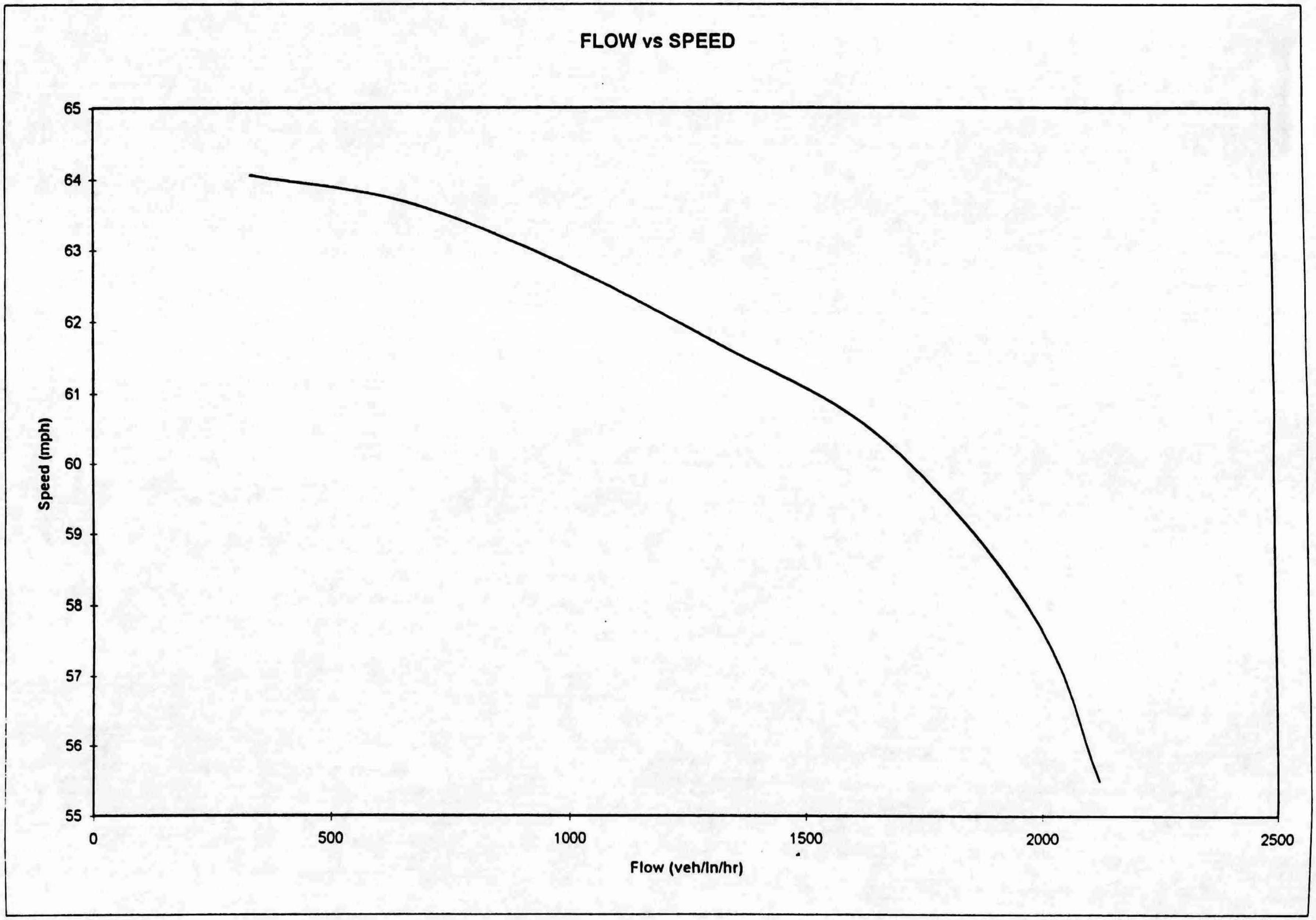

Figure 6.2 Flow vs Speed Curve.

In order to obtain the remaining part of the curve (*i.e.*, for low speeds with congested conditions), a second case was studied by introducing a 20 minute freeway incident in the simulation. This would also help to determine the freeway capacity and understand the FRESIM simulation better.

Case 2 Simulation

In the second case, for the same coded network and for a freeway volume of 5000 vph, a 20 minute incident was coded, starting at the tenth minute of the 60 minute simulation. Due to the incident, the volume would increase on the freeway, thereby dropping the speed and once the incident is over, there would be an increase in speed. The flow, speed and density values given in Table 6.2 were taken for every 5 minute intervals and the same are shown in the table below.

Table 6.2 Case 2 Simulation Results

Volume (veh/ln/hr)	**Density (Veh/ln-mile)**	**Speed (mph)**
1668	27.45	60.77
1674	27.67	60.50
1656	27.46	60.30
2040	36.12	56.48
2436	52.07	46.78
2274	61.26	37.12
1314	96.62	13.60
1200	118.69	10.11
1200	119.05	10.08
1194	118.69	10.06

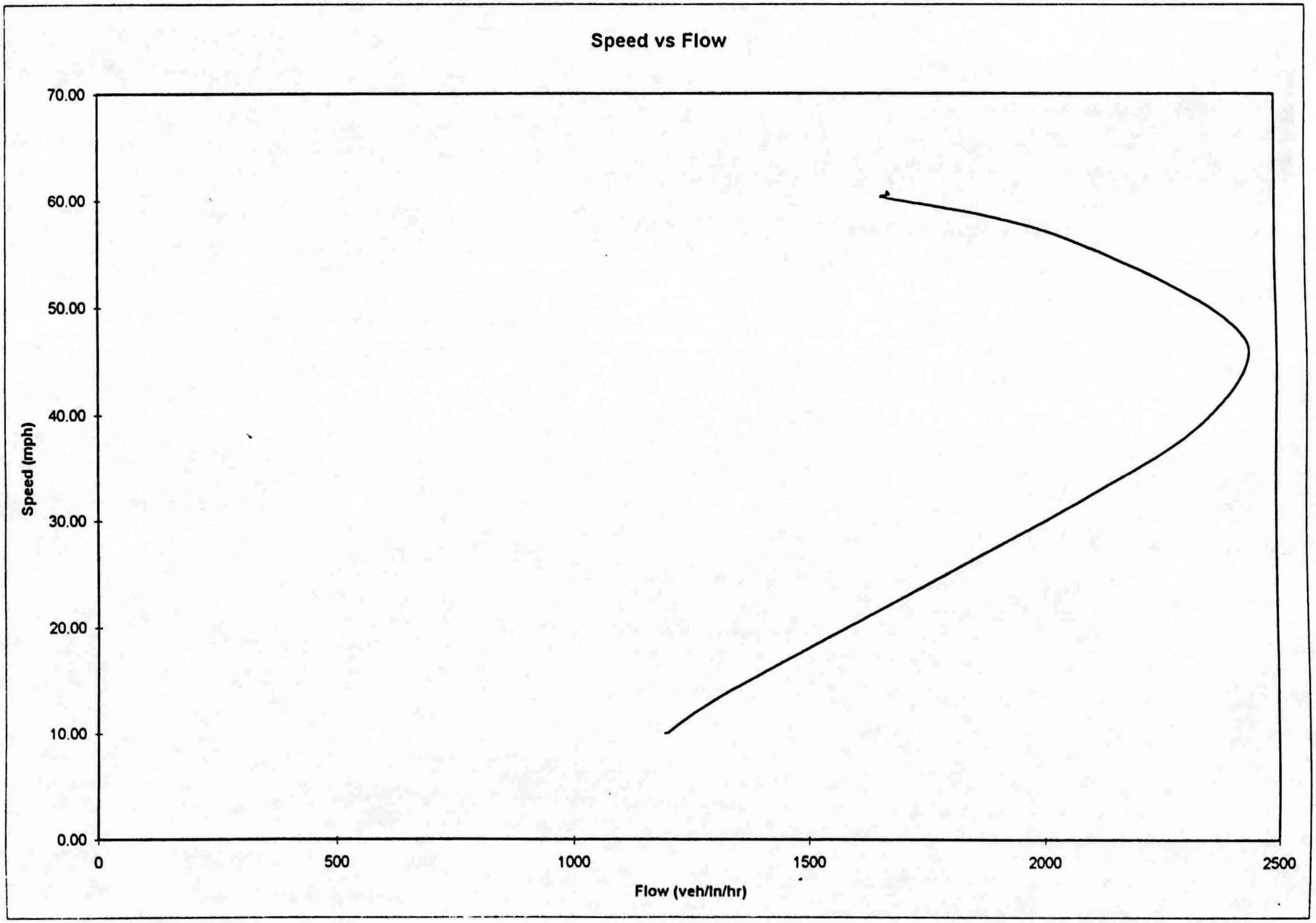

Figure 6.3 Flow vs Speed Curve.

The flow and speed values were plotted to obtain the flow-speed curve, which is shown in Figure 6.3. From the curve, it is apparent that the simulated capacity of the freeway is around 2400 veh/ln/hr. In the next case, for the same network, ramp metering was also simulated with a ramp volume of 800 vph.

Case 3 Simulation

In this case, ramp metering was simulated with a ramp volume of 800 vph and the freeway volumes ranging from 5500 vph to 6400 vph. The simulation was performed without any incident on the freeway. Out of the four ramp metering strategies that FRESIM can perform, Speed Control Metering was chosen for the simulation. For the three speed thresholds of 58 mph, 56 mph and 54 mph, headways of 6, 9 and 18 seconds were entered, respectively. The queues, total travel times and 4 headway were determined for the ramp link 121-101 and are tabulated in Table 6.3 as follows:

Table 6.3 Case 3 Simulation Results

Frwy Vol. (Vph)	Output at 1 (vph)	Input at 2 (vph)	Ramp 101-121			Link Speed of 11-1(mph)
			Headway (sec)	Queue (veh)	Total Time (sec)	
5500	6287	6281	4.74	10.3	48.9	59.04
5600	6382	6383	4.76	12.4	59.3	58.80
5700	6479	6471	4.81	18.5	89.1	56.58
5800	6562	6549	5.00	25.7	128.3	55.47
5900	6636	6615	5.09	26.5	135.2	47.43
6000	6666	6658	5.63	27.6	155.5	50.06
6100	6749	6738	5.99	28.5	170.7	40.62
6200	6754	6760	6.43	28.9	186.0	34.21
6300	6754	6760	6.54	29.4	186.0	34.21
6400	6772	6769	6.27	27.2	192.4	44.81

It was observed that for high volumes of the freeway, the queue on the entrance ramp was also high. The headway for the entrance ramp link 101-121 were calculated from the flow values and it was apparent that the increase in the headway was directly proportional to the increase in the ramp queue and was inversely proportional to the upstream freeway link speed. It was also noted that for freeway speeds less than 54 mph, the metering rate was not 18 seconds, as it should be. Therefore, further simulation studies are proposed to understand the simulation involving ramp metering. These ramp metering studies, along with other alternatives, are present issues to be further tested and resolved.

7. REFERENCES

1. Marsden, B. Ramp Meters and Travel Quality in Austin, Texas. TxDOT. May 1981.

2. Smith, S.A. INFORM Evaluation. Volume I: Executive Summary. Draft Final Report, FHWA-RD-91-075, 1992.

3. Robinson, J. and Piotrowicz, G. Ramp Metering Status in North America: Final Report. DOT-T-90-01. Federal Highway Administration, June 1995.

4. Meldrum, D.R. and Taylor, C.E. Freeway Traffic Data Prediction Using Artificial Neural Networks and Development of a Fuzzy Logic Ramp Metering Algorithm. WA-RD 365.1, Washington State Transportation Center, Seattle, Washington, April 1995.

5. Jacobson, L. Henry, K, and Mehyar, O. Real-Time Metering Algorithm for Centralized Control. Transportation Research Record #1232, 1988.

6. Nihan, N. Freeway Congestion Prediction. Draft Technical Report. Washington Transportation Center, Seattle, Washington, March 1995.

7. Papageorgiou, M., Hadj-Salem, H., and Blosseville, J.M. ALINEA: A Local Feedback Control Law for On-Ramp Metering. Transportation Research Record 1320, 1991.

8. Messer, C.J., Thomadakis, M.E., Charara, H.A., and Urbanik, T. Ramp Adaptive Metering Bottleneck Optimization (RAMBO) User's Manual. FHWA/TX/-94/1232-31. Texas Transportation Institute, November 1994.

9. Keen, K.G., Schofield, M.J., and Hay, G.C. Ramp Metering Access Control on M6 Motorway. Proc. IEE 2nd International Conf. On Road Traffic Control, London, 1986.

10. Masher, D.P., Ross, D.W., Wong, P.J., Tuan, P.L., Zeidler, H.M., and Petracek S. Guidelines for Design and Operation of Ramp Control Systems. Stanford Research Institute, Menlo Park, California, 1975.

11. Middelham, F. and Smulders, S. Isolated Ramp Metering: Real Life Study in the Netherlands. DRIVE-Office, Brussels, Belgium, 1991.

12. Hadj-Salem, H., Davee, M.M., Blosseville, J.M., and Papageorgiou, M. ALINEA: Un Outile de Regulation d'Access Isole sur Autoroute. Rapport INRETS 80, Arcueil, France, 1988.

13. Hadj-Salem, H. Blosseville, J.M., and Papageorgiou, M. ALINEA: A Local Feedback Control Law for On-Ramp Metering--A Real Life Study. Proc. 3rd IEE International Conf. On Road Traffic Control, London, 1990.

14. Zhang, H., Ritchie, S.G., and Lo, Z. A Local Neural Network Controller for Freeway Ramp Metering. Transportation Research, B. 1994.

15. Hecht-Nielsen, R. Neurocomputing, Addison-Wesley Publishing Co. 1990.

16. Greenshields, B.D. A Study of Traffic Capacity. Highway Res. Board Proc. 14, 1934.

17. JFT Associates. Freeway Simulation Model Enhancement and Integration: Tutorial Manual for FRESIM. FHWA Project DT-FH-61-85-C-00094. Federal Highway Administration., 1990.